

Integrality Gaps and Approximation Algorithms for Dispersers and Bipartite Expanders

Xue Chen^{*1}

¹Department of Computer Science, University of Texas at Austin

October 20, 2015

Abstract

We study the problem of approximating the quality of a disperser. A bipartite graph G on $([N], [M])$ is a $(\rho N, (1 - \delta)M)$ -disperser if for any subset $S \subseteq [N]$ of size ρN , the neighbor set $\Gamma(S)$ contains at least $(1 - \delta)M$ distinct vertices. Our main results are strong integrality gaps in the Lasserre hierarchy and an approximation algorithm for dispersers.

1. For any $\alpha > 0$, $\delta > 0$, and a random bipartite graph G with left degree $D = O(\log N)$, we prove that the Lasserre hierarchy cannot distinguish whether G is an $(N^\alpha, (1 - \delta)M)$ -disperser or not an $(N^{1-\alpha}, \delta M)$ -disperser.
2. For any $\rho > 0$, we prove that there exist infinitely many constants d such that the Lasserre hierarchy cannot distinguish whether a random bipartite graph G with right degree d is a $(\rho N, (1 - (1 - \rho)^d)M)$ -disperser or not a $(\rho N, (1 - \Omega(\frac{1-\rho}{\rho d+1-\rho}))M)$ -disperser. We also provide an efficient algorithm to find a subset of size exact ρN that has an approximation ratio matching the integrality gap within an extra loss of $\frac{\min\{\frac{\rho}{1-\rho}, \frac{1-\rho}{\rho}\}}{\log d}$.

Our method gives an integrality gap in the Lasserre hierarchy for bipartite expanders with left degree D . G on $([N], [M])$ is a $(\rho N, a)$ -expander if for any subset $S \subseteq [N]$ of size ρN , the neighbor set $\Gamma(S)$ contains at least $a \cdot \rho N$ distinct vertices. We prove that for any constant $\epsilon > 0$, there exist constants $\epsilon' < \epsilon$, ρ , and D such that the Lasserre hierarchy cannot distinguish whether a bipartite graph on $([N], [M])$ with left degree D is a $(\rho N, (1 - \epsilon')D)$ -expander or not a $(\rho N, (1 - \epsilon)D)$ -expander.

^{*}xchen@cs.utexas.edu

1 Introduction

In this work, we study the vertex expansion of bipartite graphs. For convenience, we always use G to denote a bipartite graph and $[N] \cup [M]$ to denote the vertex set of G . Let D and d denote the maximal degree of vertices in $[N]$ and $[M]$, respectively. For a subset S in $[N] \cup [M]$ of a bipartite graph $G = ([N], [M], E)$, we use $\Gamma(S)$ to denote its neighbor set $\{j | \exists i \in S, (i, j) \in E\}$. We consider the following two useful concepts in bipartite graphs:

Definition 1.1 A bipartite graph $G = ([N], [M], E)$ is a (k, s) -disperser if for any subset $S \subseteq [N]$ of size k , the neighbor set $\Gamma(S)$ contains at least s distinct vertices.

Definition 1.2 A bipartite graph $G = ([N], [M], E)$ is a (k, a) -expander if for any subset $S \subseteq [N]$ of size k , the neighbor set $\Gamma(S)$ contains at least $a \cdot k$ distinct vertices. It is a $(\leq K, a)$ expander if it is a (k, a) -expander for all $k \leq K$.

Because dispersers focus on hitting most vertices in $[M]$, and expanders emphasize that the expansion is in proportion of the degree D , it is often more convenient to use parameters ρ, δ , and ϵ for $k = \rho N, s = (1 - \delta)M$, and $a = (1 - \epsilon)D$ for dispersers and expanders.

These two combinatorial objects have wide applications in computer science. Dispersers are well known for obtaining non-trivial derandomization results, e.g., for derandomization of inapproximability results for MAX Clique and other NP-Complete problems [Zuc96a, TZ04, Zuc07], deterministic amplification [Sip88], and oblivious sampling [Zuc96b]. Dispersers are also closely related to other combinatorial constructions such as randomness extractors, and some constructions of dispersers follow the constructions of randomness extractors directly [TZ04, BKS⁺10, Zuc07]. Explicit constructions achieving almost optimal degree have been designed by Ta-Shma [Ta-02] and Zuckerman [Zuc07], respectively, in different important parameter regimes.

For bipartite expanders, it is well known that the probabilistic method provides very good expanders, and some applications depend on the existence of such bipartite expanders, e.g., proofs of lower bounds in different computation models [Gri01, BOT02]. Expanders also constitute an important part in other pseudorandom constructions, such as expander codes [SS96] and randomness extractors [TZ04, CRVW02, GUV09]. A beautiful application of bipartite expanders was given by Buhrman et.al. [BMRV00] in the static membership problem (see [CRVW02] for more applications and the reference therein). Explicit constructions for expansion $a = (1 - \epsilon)D$ with almost-optimal parameters have been designed in [CRVW02] and [TZ04, GUV09] for constant degree and super constant degree respectively.

We consider the natural problem of how to approximate the vertex expansion of ρN -subsets in a bipartite graph G on $[N] \cup [M]$ in terms of the degrees D, d , and the parameter ρ . More precisely, given a parameter ρ such that $k = \rho N$, it is natural to ask what is the size of the smallest neighbor set over all ρN -subsets in $[N]$. To the best of our knowledge, this question has only been studied in the context of expander graphs when G is d -regular with $M = N$ and $D = d$ by bounding the second eigenvalue. In [Kah95], Kahale proved that the second eigenvalue can be used to show the graph G is a $(\leq \rho N, \frac{D}{2})$ -expander for a $\rho \ll \text{poly}(\frac{1}{D})$. Moreover, Kahale showed that some Ramanujan graphs do not expand by more than $D/2$ among small subsets, which indicates $D/2$ is the best parameter for expanders using the eigenvalue method. In another work [WZ99], Wigderson and Zuckerman pointed out that the expander mixing lemma only helps us determine whether the bipartite graph G is a $(\rho N, (1 - \frac{4}{D\rho})N)$ -disperser or not, which is not helpful if $D\rho \leq 4$. Even if $D\rho = \Omega(1)$, the expander mixing lemma is unsatisfactory because a random bipartite graph on $[N] \cup [M]$ with right degree d is an $(N, (1 - O((1 - \rho)^d))M)$ -disperser with high probability. Therefore Wigderson and Zuckerman provided an explicit construction for the case $d\rho = \Omega(1)$ when $d = N^{1-\delta+o(1)}$ and $\rho = N^{-(1-\delta)}$.

for any $\delta \in (0, 1)$. However, there exist graphs such that the second eigenvalue is close to 1 but the graph has very good expansion property among small subsets [KV05, BGH⁺12]. Therefore the study of the eigenvalue is not enough to fully characterize the vertex expansion. On the other hand, it is well known that a random regular bipartite graph is a good disperser and a good expander simultaneously, it is therefore natural to ask how to certify a random bipartite graph is a good disperser or a good expander.

Our main results are strong integrality gaps and an approximation algorithm for the vertex expansion problem in bipartite graphs. We prove the integrality gaps in the Lasserre hierarchy, which is a strong algorithmic tool in approximation algorithm design such that most currently known semidefinite programming based algorithms can be derived by a constant number of levels in this hierarchy.

We first provide integrality gaps for dispersers in the Lasserre hierarchy. It is well known that a random bipartite graph on $[N] \cup [M]$ is an $(N^\alpha, (1 - \delta)M)$ -disperser with very high probability when N is large enough and left degree $D = \Theta_{\alpha, \delta}(\log N)$, and these dispersers have wide applications in theoretical computer science [Sha02, Zuc07]. We show an average-case complexity of the disperser problem that given a random bipartite graph, the Lasserre hierarchy cannot approximate the size of the subset in $[N]$ (equivalently the min-entropy of the disperser) required to hit at least 0.01 fraction of vertices in $[M]$ as its neighbors. The second result is an integrality gap for any constant $\rho > 0$ and random bipartite graphs with constant right degree d (the formal statements are in section 3.1).

Theorem 1.3 *(Informal Statement) For any $\alpha \in (0, 1)$ and any $\delta \in (0, 1)$, the $N^{\Omega(1)}$ -level Lasserre hierarchy cannot distinguish whether, for a random bipartite graph G on $[N] \cup [M]$ with left degree $D = O(\log N)$:*

1. G is an $(N^\alpha, (1 - \delta)M)$ -disperser,
2. G is not an $(N^{1-\alpha}, \delta M)$ -disperser.

Theorem 1.4 *(Informal Statement) For any $\rho > 0$, there exist infinitely many d such that the $\Omega(N)$ -level Lasserre hierarchy cannot distinguish whether, for a random bipartite graph G on $[N] \cup [M]$ with right degree d :*

1. G is a $(\rho N, (1 - (1 - \rho)^d)M)$ -disperser,
2. G is not a $(\rho N, (1 - C_0 \cdot \frac{1-\rho}{\rho d + 1 - \rho})M)$ -disperser for a universal constant $C_0 > 0.1$.

We also provide an approximation algorithm to find a subset of size exact ρN with a relatively small neighbor set when the graph is not a good disperser. For a balanced constant ρ like $\rho \in [1/3, 2/3]$, $\frac{\rho}{1-\rho}$ and $\frac{1-\rho}{\rho}$ are just constants, and the approximation ratio of our algorithm is close to the integrality gap in Theorem 1.4 within an extra loss of $\log d$.

Theorem 1.5 *Given a bipartite graph $([N], [M])$ that is not a $(\rho N, (1 - \Delta)M)$ -disperser with right degree d , there exists a polynomial time algorithm that returns a ρN -subset in $[N]$ with a neighbor set of size at most $(1 - \Omega(\frac{\min\{(\frac{\rho}{1-\rho})^2, 1\}}{\log d}) \cdot d(1 - \rho)^d \Delta)M$.*

For expanders, we will show that for any constant $\epsilon > 0$, there is another constant $\epsilon' < \epsilon$ such that the Lasserre hierarchy cannot distinguish the bipartite graph is a $(\rho N, (1 - \epsilon')D)$ expander or not a $(\rho N, (1 - \epsilon)D)$ expander for small ρ (the formal statement is in Section 3.2). To the best knowledge, this is the first hardness result for such an expansion property. For example, it indicates that the Lasserre hierarchy cannot distinguish between a $(\rho N, 0.6322D)$ -expander or not a $(\rho N, 0.499D)$ -expander.

Theorem 1.6 *For any $\epsilon > 0$ and $\epsilon' < \frac{e^{-2\epsilon} - (1-2\epsilon)}{2\epsilon}$, there exist constants ρ and D such that the $\Omega(N)$ -level Lasserre hierarchy cannot distinguish whether, a bipartite graph G on $[N] \cup [M]$ with left degree D :*

1. G is an $(\rho N, (1 - \epsilon')D)$ -expander;
2. G is not an $(\rho N, (1 - \epsilon)D)$ -expander.

We study the vertex expansion for a bipartite graph G on $[N] \cup [M]$ with the parameter ρ as a Constraint Satisfaction Problem (CSP) with a global constraint as follows: For $i \in [N]$, let $x_i \in \{0, 1\}$ denote whether vertex i is in the subset or not. For $j \in [M]$, j is a neighbor of the subset iff the disjunction function on j 's neighbors $\bigvee_{i \in \Gamma(j)} x_i$ is true. Then finding a ρN -subset with the fewest neighbors is the same as assigning ρN variables of $\{x_1, \dots, x_N\}$ to be 1 such that the assignment minimizes the number of satisfied constraints from $[M]$. Hence our results of Theorem 1.4 and Theorem 1.5 provide an almost tight pair of an integrality gap and an approximation algorithm for a CSP with a global constraint. We also introduce list Constraint Satisfaction Problems (list CSP) for the construction of integrality gaps for any $\rho \in (0, 1)$, which allow every variable to take k values in the alphabet instead of 1 value in the classical CSPs and relax the value of each constraint from $\{0, 1\}$ to N .

Constraint Satisfaction Problems is a class of fundamental optimization problems that has been studied in approximation algorithms and hardness of approximation for the last twenty years. For most natural CSPs, it is NP-hard to find an optimal assignment. Actually, it is even NP-hard to find an assignment that is better than a random assignment for many CSPs [Cha13]. In a surprising development, under the Unique Game Conjecture (UGC) [Kho02] tight hardness results matching integrality gaps of simple semidefinite programmings have been shown for many CSPs. Khot et.al. [KKMO07] showed dictatorship tests can be converted to UGC hardness results for CSPs. In a seminal work [Rag08], Raghavendra proved that any integrality gap of a simple semidefinite programming for a CSP can be translated to a dictatorship test with the corresponding completeness and soundness, which implies a UGC hardness result for the CSP according to [KKMO07]. Raghavendra also provided a generic algorithm for any CSP with an approximation ratio matching the integrality gap, which unifies the theory of approximation algorithms, integrality gaps, and hardness of approximation on CSPs based on UGC.

A CSP with a global constraint, which is a CSP concerning assignments restricted by an extra global cardinality constraint such as fixing the number of a given element in the assignment, is a natural generalization of CSPs but not well understood in general compared to the extensive studies in CSPs. Several important problems such as Small-Set Expansion [RS10] and Max Bisection can be formulated as a CSP with a global constraint. Small-Set Expansion hypothesis (SSE) was proposed by Raghavendra and Steurer [RS10] as a natural extension of UGC with more structures. Before stating SSE, we define the edge expansion of a subset S in a d -regular graph $H = (V, E)$ to be $\frac{E(S, V \setminus S)}{d \cdot \min\{|S|, |V \setminus S|\}}$.

Hypothesis 1.7 (*Small-Set Expansion Hypothesis [RS10]*) *For every constant $\eta > 0$, there exists a small $\delta > 0$ such that given a graph $H = (V, E)$ it is NP-hard to distinguish whether:*

1. *There exists a vertex set S of size $\delta|V|$ such that the edge expansion of S is at most η .*
2. *Every vertex sets S of size $\delta|V|$ has edge expansion at least $1 - \eta$.*

Raghavendra, Steurer and Tetali [RST10] provided an efficient algorithm that given δ and $H = (V, E)$ with edge expansion at most ϵ among subsets of size at most $\delta|V|$, it finds a subset of size $O(\delta|V|)$ with edge expansion $O(\sqrt{\epsilon \log(1/\delta)})$. In a later work, Raghavendra, Steurer and Tulsiani [RST12] proved a hardness result matching the approximation ratio for small enough ϵ that it is SSE-hard to distinguish whether the

Min Bisection of H is $O(\epsilon)$ or $\Omega(\sqrt{\epsilon})$. For other CSPs with a global constraints, even less is known. For example, Raghavendra provided a generic approximation algorithm for any integrality gaps of CSPs in [Rag08]; to the best of our knowledge, there is no known generic approximation algorithm for any CSPs with a global constraint. For Max Bisection, partition the vertex set of a graph into two parts with the same size while maximizing the crossing edges, is a natural generalization of Max-Cut problem. It is known that the approximation of Max Bisection cannot be better than Max Cut (the reduction is to make two copies of the graph), however, the best approximation ratio of Max Bisection is 0.8776 [ABG13] to our best knowledge, which is slightly smaller than the approximation ratio of Max Cut 0.8786 [GW95].

In a graph $H = (V, E)$ that is not necessarily bipartite, it is more interesting to consider the vertices in $V \setminus S$ connected to S , which is $\Gamma(S) \setminus S$, and define the vertex expansion of S to be $|V| \cdot \frac{|\Gamma(S) \setminus S|}{|S| \cdot |V \setminus S|}$ in H . Recently, Louis, Raghavendra and Vempala [LRV13] showed that vertex expansion is much harder to approximate than edge expansion, which is easy to approximate by Cheeger's inequality from the second eigenvalue. They proved that it is SSE-hard to determine whether the vertex expansion of a given graph is at most $O(\epsilon)$ or at least $\Omega(\sqrt{\epsilon \log d})$ for small enough ϵ . At the same time, they also provided an efficient algorithm based on semidefinite programmings with an asymptotic matching approximation ratio that given a graph with vertex expansion ϵ and bounded degree d , finds a subset with vertex expansion $O(\sqrt{\epsilon \log d})$.

When the vertex expansion in expanders is independent of the left degree, we prove that it is SSE-hard to distinguish between good expanders and bad expanders when ρ is small enough and degree is large enough by amplifying the gap in the hardness result of [LRV13](see Theorem 5.3 for a formal statement).

Theorem 1.8 *(Informal Statement) For any small constant δ and any constant $\Delta > 1 + \delta$, given a bipartite graph G on $[N] \cup [M]$ with ρ small enough and left degree D large enough, it is SSE-hard to distinguish between*

1. *There exists a ρN subset of $[N]$ with at most $(1 + \delta) \cdot \rho N$ neighbors.*
2. *Every ρN subset of $[N]$ has at least $\Delta \cdot \rho N$ neighbors.*

In another extreme case that the bipartite graph G has a ρN -subset with at most $(1 + \epsilon)\rho M$ neighbor, We provide an efficient algorithm with an asymptotic matching approximation ratio by following the previous work of [LRV13, CMM06, BFK⁺11, LM14].

Theorem 1.9 *(Informal Statement) Given a regular bipartite graph on $[N] \cup [M]$ that is D -regular in $[M]$ and d -regular in $[N]$, suppose $d|D$ and the smallest neighbor set of ρN -subsets in $[N]$ is $\rho(1 + \epsilon)M$. There is a polynomial time algorithm that finds a subset S of size $[0.99\rho N, 1.01\rho N]$ with at most $(1 + \tilde{O}_\rho(\sqrt{\epsilon \log d / \rho}))|S| \frac{M}{N}$ neighbors.*

This paper is organized as follows. We will define some basic notations and provide some background for our problems, then we give a brief overview of our proof in Section 2. We prove the integrality gaps of Theorem 1.6, Theorem 1.3, and Theorem 1.4 in Section 3, and provide the approximation algorithm of Theorem 1.5 in Section 4. For bipartite graphs with a ρN -subset of at most $(1 + \epsilon)\rho M$ neighbors, we prove the hardness result and provide the approximation algorithm in Section 5.

1.1 Discussion

We study the vertex expansion of bipartite graphs as a list CSP with a global constraint and provide an integrality gap in Theorem 1.4 and an approximation algorithm in Theorem 1.5 that are almost tight to each other. It is therefore of great interest to prove a hardness result matching the integrality gap and the

approximation ratio. Not only will this unify integrality gaps, hardness of approximation and approximation algorithms for CSPs with a global constraint, but also it will provide an explicit construction of a $(\rho N, 1 - (1 - \rho)^d M)$ -disperser, which beats all known constructions of dispersers and matches the parameters from the probabilistic method. The construction is simple: suppose there is a reduction from SSE (UGC) to the disperser problem with d and ρ . Start with a known instance in the sound case of SSE (UGC) from [KV05, BGH⁺12] and follow the reduction to obtain a bipartite graph G on $[N] \cup [M]$. G is in the sound case from the property of the reduction, which demonstrates it is a $(\rho N, (1 - (1 - \rho)^d)M)$ -disperser.

It is known that UGC is not enough to prove a hardness result for vertex expansion or edge expansion [RS10, RST12], hence it is interesting to further investigate the Small Set Expansion Hypothesis. More precisely, a common way to prove the hardness of a CSP is to construct a dictatorship test corresponding to the CSP. The dictatorship test corresponding to the vertex expansion problem with completeness $1 - \frac{1-\rho}{\rho d+1-\rho}$ and soundness $1 - (1 - \rho)^d$ for infinitely many d is known from [BGGP12, DM13]. The standard reduction from UGC to CSPs using dictatorship tests [KKMO07] always apply folding to balance each boolean cube. However, folding operation (negation) is not supported in expansion problems.

To the best of our knowledge, all known reductions [RST12, LRV13] from the Small Set Expansion hypothesis only work for dictatorship tests with small noise, but the dictatorship test mentioned above with completeness $1 - \frac{1-\rho}{\rho d+1-\rho}$ and soundness $1 - (1 - \rho)^d$ requires pairwise independence. Because such a reduction from SSE to the disperser problem would provide an explicit construction matching the construction from the probabilistic method, it is interesting to discover more reductions from the Small Set Expansion hypothesis that support more dictatorship tests, which include tests with pairwise independence. Although we show a hardness result for the vertex expansion in bipartite graphs from previous work [LRV13] based on SSE, the hardness result does not shed any light on the relations between ρ , the left and right degrees, and the expansion in the bipartite graph. Because the hardness result in [LRV13] only works for very small ϵ like $< 10^{-10}$, the parameters ρ, D become exponential in ϵ after amplification.

It is of great interest to further study the hardness and integrality gaps of (k, A) -expanders in terms of the left degree D and ρ . Observe that the integrality gap of Theorem 1.6 in terms of ρ and d matches the soundness and the completeness in the dictatorship test of [BGGP12] for parameters $\rho d < 1$. However, our estimation fails to provide an integrality gap for $(\rho N, D - 1.1)$ -expanders even for $(\rho N, D - \sqrt{D})$ -expanders. It is well known that a balanced random bipartite graph ($N = \Theta(M)$ and $D = \Theta(d)$) is a $(\rho N, D - 1.1)$ -expander for $\rho \leq \exp(-D)$ with high probability, and such a probabilistic construction plays an important role in the proofs of different lower bounds [Gri01, Tul09, BOT02]. Our estimation in the Lasserre hierarchy for ρN -subsets is $(1 - \epsilon)D \cdot \rho N$ for $\epsilon = O(\rho d)$. Because $\epsilon = O(\rho d) < 1/D$ for $\rho \leq \exp(-D)$, the estimation does not provide a meaningful integrality gap any more. It is natural to ask what is the integrality gap of $(\rho N, D - 1.1)$ -expanders and what is the integrality gap in terms of D and ρ ; the first problem was already asked by Barak [Bar14]. Therefore it would be interesting to scale down the degree D in the integrality gap and show more integrality gaps of expanders in terms of ρ and D .

It is interesting to design an algorithm with an approximation ratio matching the integrality gaps especially for integrality gap in Theorem 1.6 and Theorem 1.4. Such an algorithm may have other applications in computer science like generating a random bipartite graph and verifying that it is a good expander/disperser. At the same time, it is even of great interest to provide a generic approximation algorithm matching the integrality gap and hardness for any CSP with a global constraint.

2 Preliminaries

For simplicity, we assume the bipartite graph is d -regular on $[M]$. The expected number of neighbors of a random subset $S \subseteq [N]$ of size ρN is $E[|\Gamma(S)|] = M \cdot \left(1 - \frac{N-|S|}{N} \cdot \frac{N-1-|S|}{N-1} \dots \frac{(N-d+1)-|S|}{N-d+1}\right) = (1 - (1 - \rho)^d + o(1))M$, which demonstrates there is a subset S of size ρN with $|\Gamma(S)|$ at most $(1 - (1 - \rho)^d + o(1))M$. On the other hand, $|\Gamma(S)|/M \geq \rho$ for any subset S of size ρN if the bipartite graph is D -regular bipartite on $[N]$. Sometime, it is more convenient to work with δ on a $(\rho N, (1 - \delta)M)$ -disperser and the approximation ratio of the algorithm in Section 4 is in terms of δ .

For convenience, let $\binom{[n]}{k}$ denote the subsets in $[n]$ with size k and $\binom{[n]}{\leq k}$ denote the subsets of size at most k . We always use 1_E to denote the indicator function of an event E , e.g., $1_{x_i=1} = x_i$ for $x_i \in \{0, 1\}$. We say $C \subseteq F_q^d$ is a pairwise independent subspace of F_q^d iff C is a subspace and any 2 variables in the uniform distribution of C are independent.

In this work, we always use Λ to denote a Constraint Satisfaction Problem and Φ to denote an instance of the CSP. A CSP Λ is specified by a width d , a finite field F_q for a prime power q and a predicate $C \subset F_q^d$. An instance Φ of Λ consists of n variables and m constraints such that every constraint j is in the form of $x_{j,1} \cdots x_{j,d} \in C + \vec{b}_j$ for some $\vec{b}_j \in F_q^d$ and d variables $x_{j,1}, \dots, x_{j,d}$.

We prove our integrality gaps in the Lasserre hierarchy. It is a variant of the hierarchies that have been studied by several authors including Shor [Sho87], Parrilo [Par03], Nesterov [Nes00] and Lasserre [Las02]. For convenience, we adopt to the notations of the Lasserre hierarchy and provide a description of the Lasserre hierarchy in Section 2.2.

2.1 Proof Overview

We outline our approaches in this section.

To prove the integrality gaps of vertex expansion in the Lasserre hierarchy, we first illustrate the idea to prove the integrality gaps of dispersers and then move to the integrality gaps of expanders. We start with a random graph G that is d -regular on the right such that it is a $(\rho N, (1 - (1 - \rho)^d - o(1))M)$ -disperser from $[N]$ to $[M]$ and a $(\leq \exp(-d)M, d - 1.1)$ -expander from $[M]$ to $[N]$ (this happens with high probability). Next we write a natural $\{0, 1\}$ -programming of vertex expansion among ρN -subsets in G as a CSP with a global constraint $\sum x_i = \rho N$ and an objective function $\min \sum_{j \in [M]} 1_{\vee_{i \in \Gamma(j)} x_i}$, which seeks the size of the smallest neighbor set over all ρN -subsets in $[N]$. For convenience, we rewrite the objective function as $\min \sum_{j \in [M]} (1 - 1_{\wedge_{i \in \Gamma(j)} \bar{x}_i}) = M - \max_{j \in [M]} 1_{\wedge_{i \in \Gamma(j)} \bar{x}_i}$ such that it looks like a standard CSP of width d that maximizes the objective value.

Now let us turn to the SDP solution in the Lasserre hierarchy for vertex expansion among ρN -subsets. A first try would be to think each vertex $i \in [N]$ corresponds to a variable in the CSP, and each vertex in M corresponds to a constraint $\wedge_{i \in \Gamma(j)} \bar{x}_i$. It is a $(\leq \exp(-d)M, d - 1.1)$ -expander from the right hand side $[M]$ to the left hand side $[N]$ such that it has very good variable expansion property in constraints. Hence it is possible to use the known construction of SDP solution in the Lasserre hierarchy on random instances of MAX-CSPs in [Gri01, Sch08, Tul09, Cha13]. However, this approach has one problem and only works for $\rho = 1 - 1/q$.

The first thing is to verify the SDP solution satisfies the global constraint $\sum \bar{x}_i = (1 - \rho)N$, namely the matrix induced by the global constraint is positive semidefinite. An important ingredient in our proof is from the work of Guruswami, Sinop and Zhou [GSZ14], which proves that the matrix induced by the global constraint is positive semidefinite as long as the summation of vectors satisfy the constraint $\sum \vec{v}_i = \rho N \cdot \vec{v}_\emptyset$. Actually, it is not difficult to prove $\sum \vec{v}_i = \rho N \cdot \vec{v}_\emptyset$ is also necessary in the vertex expansion problem because of the equation $\sum x_i = \rho N$. To obtain $\sum \vec{v}_i = \rho N \cdot \vec{v}_\emptyset$, we assume that the number of variables in the CSP

is n such that $[N] = [n] \times F_q$ instead of N variables and each vertex in $[N]$ corresponds to a variable with a label in F_q and $\sum_{i \in [n] \times F_q} \vec{v}_i = n \cdot \vec{v}_\emptyset$. We notice that such a reduction also works for the SDP solution in the Lasserre hierarchy and provide a estimation $(1 - \frac{1}{(q-1)d+1})M$ for the small neighbor set among ρN -subsets for $\rho = 1 - 1/q$ given F_q as the alphabet of the CSP. $\rho = 1 - 1/q$ comes from the fact that the objective function is in terms of \bar{x}_i and $1/q$ fraction of variables are true in the SDP solution of MAX-CSPs.

To generalize the integrality gap for any $\rho = 1 - k/q$ especially for $\rho = 1/q$ and $k = q - 1$, we introduce list Constraint Satisfaction Problems that allow each variable x_i to take k values in the alphabet and relax the value of one constraint from $\{0, 1\}$ to N^+ . Our main technical lemma is to prove an lower bound on the SDP value of list CSPs in the Lasserre hierarchy such that we could provide an integrality gap of vertex expansion for any ρ . The method introduced by Grigoriev [Gri01] and rediscovered by Schoenebeck [Sch08] for CSPs using resolution proofs does not work for list CSPs, because the resolution proofs become difficult when each variable is allowed to take k values. Instead of following the previous method, we study an extra property about the pairwise independent predicate $C \subseteq F_q^d$, which tries to find a k -subset Q in the alphabet maximizing $|Q^d \cap C|$. Then we utilize the SDP solution from standard CSP and redefine $x_i = \alpha + Q$ in the list CSP if $x_i = \alpha$ in the CSP. The rest of the proof is to work out the SDP solution in the Lasserre hierarchy and verify the equation $\sum_i \vec{v}_i = \rho N \vec{v}_\emptyset$ in order to satisfy the global constraint. More detail of the pairwise independent subspace can be found in Section 2.3, the formal definition of list CSP and the estimation of SDP value of list CSPs in the Lasserre hierarchy can be found in Section 3. Thus we could obtain an integrality gap for the disperser problems with any ρ .

To obtain the integrality gap of expander problems, we notice the estimation of SDP value for ρN -subsets in the Lasserre hierarchy is at most $(\rho d - \rho^2 \binom{d}{2})M$ when $\rho d < 1$. If the bipartite graph is also D -regular in $[N]$, we rewrite it as $(1 - \rho(d-1)/2)\rho d M = (1 - \rho(d-1)/2)D \cdot \rho N$ from the equation $dM = DN$. Therefore we get an estimation of $(1 - \epsilon)D$ expansion for ρN -subsets in the Lasserre hierarchy. So the proof of Theorem 1.6 is to follow the above proof of general ρ and generate a bipartite graph with the integrality gap that is almost D -regular in $[N]$ and almost d -regular in $[M]$.

Our approximation algorithm for a $(\rho N, (1 - \Delta)M)$ -disperser follows the approach of Hast [Has05] and Charikar et.al. [CMM07] by choosing a deliberate preprocessor. The integrality gap implies that the approximation ratio on Δ can be at most $O(\frac{\rho d + 1 - \rho}{1 - \rho} \cdot (1 - \rho)^d)$ in the Lasserre hierarchy. We extend the analysis of [AN04] and [CMM07] to achieve an approximation ratio of $\Omega(\frac{\min\{(\frac{\rho}{1-\rho})^2, 1\}}{\log d} \cdot d \cdot (1 - \rho)^d)$. The main difficulty of the algorithm is to guarantee that the size of the subset returned is exact ρN . Otherwise, for $\rho = 1/2$, a random $(1 - \frac{1}{\sqrt{d}}) \cdot \frac{1}{2}N$ -subset can guarantee a neighbor set of size $\leq (1 - (\frac{1}{2} + \frac{1}{\sqrt{2d}}))^d M$ that beats the integrality gap if d is large enough, because the size of the subset is smaller than the target. One common method to round the SDP of a CSP is to take the inner product of every vector in the SDP and a Gaussian vector as a real value for every variable [GW95, MM12]. However, there is less known about how to satisfy the global constraint.

We first generalize the rounding algorithm of [CMM07] to guarantee a subset of size $(1 \pm \frac{1}{\sqrt{d}})\rho N$. To further obtain a subset of size exact ρN , let \vec{v}_i denote the vector for each vertex i in the left hand side. We insist on adding a new constraint $\sum_{i \in [N]} \vec{v}_i = \vec{0}$ in the SDP to bound the size of the subset because it provides an extra property $\sum_{i \in [N]} \langle \vec{v}_i, \vec{g} \rangle = 0$ for any vector \vec{g} . Then we replace the first step in the algorithm of [CMM07] by a more cautious rounding process, which is motivated by the work of Alon and Noar [AN04]. Eventually, our algorithm guarantees a ρN -subset with an extra loss of $\frac{\min\{\frac{\rho}{1-\rho}, \frac{1-\rho}{\rho}\}}{\log d}$ on the approximation ratio compared to the integrality gap, whose loss $\log d$ is from the preprocessor and $\min\{\frac{\rho}{1-\rho}, \frac{1-\rho}{\rho}\}$ is from the constraints $\sum_{i \in [N]} \vec{v}_i = \vec{0}$ and $|\vec{v}_i| \leq 1$ for optimal solution. However, our algorithm do not work for expanders because of the constant lost in front of Δ .

For $(\rho N, \rho(1 + \epsilon)M)$ -dispersers, the hardness result is based on the work of Louis, Raghavendra and Vempala [LRV13], which provide a basic hardness result like $1 + \epsilon$ and $1 + \sqrt{\epsilon \log d}$. Then we amplify the gap using graph products by enlarging the degrees in the bipartite graph. For the approximation algorithm, we follow the approach of Louis and Makarychev [LM14] and apply the idea of finding balanced cut by the sparsest cut algorithm because the expansion of the graph is very small in this case.

2.2 Lasserre Hierarchy

We provide a short description the semidefinite programming relaxations from the Lasserre hierarchy [Las02] (see [Rot13, Bar14] for a complete introduction of Lasserre hierarchy and sum of squares proofs). We will use $f \in \{0, 1\}^S$ for $S \subset [N]$ to denote an assignment on variables in $\{x_i | i \in S\}$. Conversely, let f_S denote the partial assignment on S for $f \in \{0, 1\}^n$ and $S \subset [n]$. For two assignments $f \in \{0, 1\}^S$ and $g \in \{0, 1\}^T$ we use $f \circ g$ to denote the assignment on $S \cup T$ when f and g agree on $S \cap T$. For a matrix A , we will use $A_{(i,j)}$ to describe the entry (i, j) of A and $A \succeq 0$ to denote that A is positive semidefinite.

Consider a $\{0, 1\}$ -programming with an objective function Q and constraints P_0, \dots, P_m , where Q, P_0, \dots, P_m are from $\binom{[n]}{\leq d} \times \{0, 1\}^d$ to \mathbb{R} :

$$\begin{aligned} & \max \sum_{R \subset \binom{[n]}{\leq d}, h \in \{0, 1\}^R} Q(R, h) 1_{x_R=h} \\ \text{Subject to } & \sum_{R \subset \binom{[n]}{\leq d}, h \in \{0, 1\}^R} P_j(R, h) 1_{x_R=h} \geq 0 & \forall j \in [m] \\ & x_i \in \{0, 1\} & \forall i \in [n] \end{aligned}$$

Let $y_S(f)$ denote the probability that the assignment on S is f in the pseudo-distribution. This $\{0, 1\}$ -programming [Las02] in the t -level Lasserre hierarchy is:

$$\begin{aligned} & \max \sum_{R \subset \binom{[n]}{\leq d}, h \in \{0, 1\}^R} Q(R, h) y_R(h) \\ \text{Subject to } & (y_{S \cup T}(f \circ g))_{(S \subset \binom{[n]}{\leq t}, f \in \{0, 1\}^S), (T \subset \binom{[n]}{\leq t}, g \in \{0, 1\}^T)} \succeq 0 \\ & \left(\sum_{R \subset \binom{[n]}{\leq d}, h \in \{0, 1\}^R} P_j(R, h) y_{S \cup T \cup R}(f \circ g \circ h) \right)_{(S \subset \binom{[n]}{\leq t}, f \in \{0, 1\}^S), (T \subset \binom{[n]}{\leq t}, g \in \{0, 1\}^T)} \succeq 0, \forall j \in [m] \end{aligned} \tag{1}$$

$$\tag{2}$$

An important tool in the Lasserre hierarchy to prove that the matrices in (2) are positive semidefinite is introduced by Guruswami, Sinop and Zhou in [GSZ14], we restate it here and prove it for completeness. Let $u_S(f)$ for all $S \in \binom{[n]}{\leq t}, f \in \{0, 1\}^S$ be the vectors to explain the matrix (1).

Lemma 2.1 (Restatement of Theorem 2.2 in [GSZ14]) *If $\sum_{R,h} P(R, h) \vec{u}_R(h) = \vec{0}$, then the corresponding matrix in (2) is positive semidefinite.*

Proof. From the definition of \vec{u} , we have

$$\begin{aligned} \sum_{R,h} P(R, h) y_{S \cup T \cup R}(f \circ g \circ h) &= \left\langle \sum_{R,h} P(R, h) \vec{u}_{S \cup R}(f \circ h), \vec{u}_T(g) \right\rangle \\ &= \langle \vec{u}_{S \cup T}(f \circ g), \sum_{R,h} P(R, h) \vec{u}_R(h) \rangle = 0. \end{aligned}$$

□

2.3 Subspace

We introduce an extra property of pairwise independent subspaces for our construction of integrality gaps of list Constraint Satisfaction Problems.

Definition 2.2 Let C be a pairwise independent subspace of F_q^d and Q be a subset of F_q with size k . We say that C stays in Q with probability p if $\Pr_{x \sim C}[x \in Q^d] = \frac{|C \cap Q^d|}{|C|} \geq p$.

In [BGGP12], Benamini et.al. proved $p \leq \frac{k/q}{(1-k/q) \cdot d + k/q}$ for infinitely many d when $|Q| = k$. They also provided a distribution that matches the upper bound with probability $\frac{k/q}{(1-k/q) \cdot d + k/q}$ for every d with $q|(d-1)$. In our work, we need the property that C is a subspace rather than an arbitrary distribution in F_q^d . We provide two constructions for the base cases $k = 1$ and $k = q - 1$.

Lemma 2.3 There exist infinitely many d such that there is a pairwise independent subspace $C \subset F_q^d$ that stays in a size 1 subset Q of F_q with probability $\frac{1/q}{(1-1/q)d+1/q}$.

Proof. Choose $Q = \{0\}$ and C to be the dual code of Hamming codes over F_q with block length $d = \frac{q^l-1}{q-1}$ and distance 3 for an integer l . Using $|C| = q^l$, the probability is $\frac{1}{|C|} = \frac{1/q}{(1-1/q)d+1/q}$. It is pairwise independent because the dual distance of C is 3. □

Lemma 2.4 There exist infinitely many d such that there is a pairwise independent subspace $C \subset F_q^d$ staying in a $(q-1)$ -subset Q of F_q with probability at least $\Omega(\frac{(q-1)/q}{d/q+(q-1)/q})$.

Proof. First, we provide a construction for $d = q - 1$ then generalize it to $d = (q - 1)q^l$ for any integer l . For $d = q - 1$, the generator matrix of the subspace is a $(q - 1) \times 2$ matrix where row i is (α_i, α_i^2) for $q - 1$ distinct elements $\{\alpha_1, \dots, \alpha_{q-1}\} = F_q^*$. Because $\alpha_i \neq \alpha_j$ for any two different rows i and j , it is pairwise independent. Let $Q = F_q \setminus \{1\}$. Using the inclusion-exclusion principle and the fact that a quadratic equation can have at most 2 roots in F_q :

$$\begin{aligned} \Pr_{x \sim C}[x \in Q^d] &= \Pr_{x \sim C}[\forall \beta \in F_q^*, x_\beta \neq 1] \\ &= 1 - \sum_{\beta \in F_q^*} \Pr[x_\beta = 1] + \sum_{\{\beta_1, \beta_2\} \in \binom{F_q^*}{2}} \Pr[x_{\beta_1} = 1 \wedge x_{\beta_2} = 1] \\ &\quad - \sum_{\{\beta_1, \beta_2, \beta_3\} \in \binom{F_q^*}{3}} \Pr[x_{\beta_1} = 1 \wedge x_{\beta_2} = 1 \wedge x_{\beta_3} = 1] + \dots \\ &= 1 - \frac{q-1}{q} + \frac{\binom{q-1}{2}}{q^2} - 0 + 0 \\ &= \frac{q^2 - q + 2}{2q^2} \geq \frac{1}{2} - \frac{1}{2q} \end{aligned}$$

For any $d = (q - 1)q^l$, the generator matrix of the subspace is a $d \times (l + 2)$ matrix where every row is in the form $(\alpha, \alpha^2, \beta_1, \dots, \beta_l)$ for all nonzero elements $\alpha \in F_q^*$ and $\beta_1 \in F_q, \dots, \beta_l \in F_q$. The pairwise independence comes from a similar analysis. $\Pr_{x \sim C}[x \in Q^d] \geq \frac{1}{q^l}(\frac{1}{2} - \frac{1}{2q})$ because it is as same as $d = q - 1$ when all coefficients before β_1, \dots, β_l are 0, which is $\geq \frac{1}{3} \cdot \frac{q-1}{d+q-1}$. □

Remark 2.5 The construction for $d = q - 1$ also provides a subspace that stays in Q with probability $1 - \frac{d}{q} + \frac{\binom{d}{2}}{q^2}$ for any $d < q - 1$ by deleting unnecessary rows in the matrix.

3 Integrality Gaps

We first consider a natural $\{0, 1\}$ programming to determine the vertex expansion of ρN -subsets in $[N]$ given a bipartite graph $G = ([N], [M], E)$:

$$\begin{aligned} \min \sum_{j=1}^M \mathbb{V}_{i \in \Gamma(j)} x_i &= \min \sum_{j=1}^M (1 - \mathbb{1}_{\mathbb{V}_{i \in \Gamma(j), x_i=0}}) \\ \text{Subject to } \sum_{i=1}^N x_i &\geq \rho N \\ x_i &\in \{0, 1\} \quad \text{for every } i \in [N] \end{aligned}$$

We relax it to a convex programming in the t -level Lasserre hierarchy.

$$\begin{aligned} \min \sum_{j=1}^M (1 - y_{\Gamma(j)}(\vec{0})) \\ \text{Subject to } (y_{S \cup T}(f \circ g))_{((S \in \binom{[N]}{\leq t}, f \in \{0, 1\}^S), (T \in \binom{[N]}{\leq t}, g \in \{0, 1\}^T))} &\succeq 0 \quad (3) \\ \left(\sum_{i=1}^N y_{S \cup T \cup \{i\}}(f \circ g \circ 1) - \rho N \cdot y_{S \cup T}(f \circ g) \right)_{((S \in \binom{[N]}{\leq t}, f \in \{0, 1\}^S), (T \in \binom{[N]}{\leq t}, g \in \{0, 1\}^T))} &\succeq 0 \quad (4) \end{aligned}$$

In this section, we focus on random bipartite graphs G on $[N] \cup [M]$ that are d -regular in $[M]$, which are generated by connecting each vertex in $[M]$ to d random vertices in $[N]$ independently. The main technical result we will prove in this section is:

Lemma 3.1 Suppose there is a pairwise independent subspace $C \subseteq F_q^d$ staying in a k -subset with probability $\geq p_0$. Let $G = ([N], [M], E)$ be a random bipartite graph with $M = O(N)$ that is d -regular in $[M]$, the $\Omega(N)$ -level Lasserre hierarchy for G and $\rho = 1 - k/q$ has an objective value at most $(1 - p_0 + \frac{1}{N^{1/3}})M$ with high probability.

We introduce list Constraint Satisfaction Problems which allow every variable to take k values from the alphabet. Next, we lower bound the objective value of an instance of a list CSP in the Lasserre hierarchy from the objective value of the corresponding instance of the CSP in the Lasserre hierarchy. Then we show how to use list CSPs to obtain an upper bound of the vertex expansion for $\rho = 1 - k/q$ in the Lasserre hierarchy.

Definition 3.2 (list Constraint Satisfaction Problem) A list Constraint Satisfaction Problem (list CSP) Λ is specified by a constant k , a width d , a domain over finite field F_q for a prime power q , and a predicate $C \subseteq F_q^d$. An instance Φ of Λ consists of a set of variables $\{x_1, \dots, x_n\}$ and a set of constraints $\{C_1, C_2, \dots, C_m\}$ on the variables. Every variable x_i takes k values in F_q , and every constraint C_j consists of a set of d variables $x_{j,1}, x_{j,2}, \dots, x_{j,d}$ and an assignment $\vec{b}_j \in F_q^d$. The value of C_j is $|(C + \vec{b}_j) \cap x_{i,1} \times x_{i,2} \times \dots \times x_{i,d}| \in \mathbb{N}$. The value of Φ is the summation of values over all constraints, and the objective is to find an assignment on $\{x_1, \dots, x_n\}$ that maximizes the total value as large as possible.

Remark 3.3 We abuse the notation C_j to denote the variable subset $\{x_{j,1}, x_{j,2}, \dots, x_{j,d}\}$. Our definition is consistent with the definition of the classical CSP when $k = 1$. The differences between a list CSP and a classical CSP are that a list CSP allow each variable to choose k values in F_q instead of one value and relax every constraint C_i from $F_q^d \rightarrow \{0, 1\}$ to $F_q^d \rightarrow \mathbb{N}$.

The $\{0, 1\}$ programming for an instance Φ with variables $\{x_1, \dots, x_n\}$ and constraints $\{C_1, \dots, C_m\}$ of Λ with parameters k, F_q , and a predicate C states as follows (the variable set is the direct product of $[n]$ and F_q in the $\{0, 1\}$ programming):

$$\begin{aligned} & \max \sum_{j \in [m]} \sum_{f \in C + \vec{b}_j} 1_{\forall i \in C(j), x_{i,f(i)}=1} \\ & \text{Subject to } x_{i,\alpha} \in \{0, 1\} \quad \forall (i, \alpha) \in [n] \times F_q \\ & \sum_{\alpha \in F_q} x_{i,\alpha} = k \quad \forall i \in [n] \end{aligned}$$

The SDP in the t -level Lasserre hierarchy for Φ succeeds this $\{0, 1\}$ programming as follows:

$$\begin{aligned} & \max \sum_{j \in [m]} \sum_{f \in C + \vec{b}_j} y_{(C_j, f)}(\vec{1}) \\ & \text{S.t. } (y_{S \cup T}(f \circ g))_{(S \subset \binom{[n] \times F_q}{\leq t}, f \in \{0, 1\}^S), (T \subset \binom{[n] \times F_q}{\leq t}, g \in \{0, 1\}^T)} \succeq 0 \end{aligned} \quad (5)$$

$$(k \cdot y_{S \cup T}(f \circ g) - \sum_{\alpha} y_{S \cup T \cup \{(i, \alpha)\}}(f \circ g \circ 1))_{(S \subset \binom{[n] \times F_q}{\leq t}, f \in \{0, 1\}^S), (T \subset \binom{[n] \times F_q}{\leq t}, g \in \{0, 1\}^T)} = 0, \forall i \in [n] \quad (6)$$

Definition 3.4 Let Λ be the list CSP problem with parameters k, q, d and a predicate $C \subset F_q^d$. Let Φ be an instance of Λ with n variables and m constraints. $p(\Phi)$ is the projection instance from Φ in the CSP of the same parameters $q, d, C \subseteq F_q^d$, and the same constraints $(C_1, \vec{b}_1), (C_2, \vec{b}_2), \dots, (C_m, \vec{b}_m)$ except $k = 1$.

Recall that a subspace $C \subset F_q^d$ stays in a subset $Q \subset F_q$ with probability p_0 if $\Pr_{x \sim C}[x \in Q^d] \geq p_0$. We lower bound Φ 's objective value in the Lasserre hierarchy by exploiting the subspace property of C and Q .

Lemma 3.5 Let Φ be an instance of the list CSP Λ with parameters k, q, d and a predicate C , where C is a subspace of F_q^d staying in a k -subset Q with probability at least p_0 . Suppose $p(\Phi)$'s value is γ in the w -level Lasserre hierarchy, then Φ 's value is at least $p_0|C| \cdot \gamma$ in the w -level Lasserre hierarchy.

Proof. Let $y_S(f)$ and $\vec{v}_S(f)$ for $S \in \binom{[n] \times F_q}{\leq w}$ and $f \in \{0, 1\}^S$ denote the pseudo-distribution and the vectors in the w -level Lasserre hierarchy for $p(\Phi)$ respectively. Let z and \vec{u} denote the pseudo-distribution and vectors in the w -level Lasserre hierarchy for Φ . The construction of z and \vec{u} from y and \vec{v} are based on the subspace C and Q . The intuition is to choose $x_i = \alpha + Q$ in Φ if $x_i = \alpha$ for some $\alpha \in F_q$ in $p(\Phi)$.

Before constructing z and \vec{u} , define \oplus operation as follows. For any $S \in \binom{[n] \times F_q}{\leq w}$, $g \in \{0, 1\}^S$, and $P \subseteq F_q$, let $S \oplus P$ denote the union of the subset $(i, \alpha + P)$ for every element (i, α) in S , which is $\cup_{(i, \alpha) \in S} \{(i, \alpha + P)\}$ in $[n] \times F_q$, and $g \oplus P \in \{0, 1\}^{S \oplus P}$ denote the assignment on $S \oplus P$ such that $g \oplus P(i, \alpha + P) = g(i, \alpha)$. If there is a conflict in the definition of $g \oplus P$, namely $\exists (i, \beta)$ such that $(i, \beta) \in (i, \alpha_1 + P)$ and $(i, \beta) \in (i, \alpha_2 + P)$ for two distinct $(i, \alpha_1), (i, \alpha_2)$ in S , define $g \oplus P(i, \beta)$ to be an arbitrary one. Because every variable only takes one value in $p(\Phi)$, $y_S(g) = 0$ if there is a conflict

on $g \oplus P \in \{0, 1\}^{S \oplus P}$. Follow the intuition mentioned above, for any $S \subset \binom{[n] \times F_q}{\leq w}$ and $g \in \{0, 1\}^S$, let $R = \{i | \exists \alpha, (i, \alpha) \in S\}$,

$$\begin{aligned} z_S(g) &= \sum_{T \in \binom{R \times F_q}{\leq w}, g' \in \{0, 1\}^T : S \subseteq T \oplus Q, g' \oplus Q(S) = g} y_T(g'), \\ \vec{u}_S(g) &= \sum_{T \in \binom{R \times F_q}{\leq w}, g' \in \{0, 1\}^T : S \subseteq T \oplus Q, g' \oplus Q(S) = g} \vec{v}_T(g'). \end{aligned}$$

The verification of the fact that \vec{u} explains z in (5) of Φ is straightforward. To verify (6) is positive semidefinite, notice that every variable x_i takes k values in F_q :

$$\sum_{\alpha \in F_q} z_{(i, \alpha)}(1) = \sum_{\alpha \in F_q} \sum_{\beta \in Q} y_{(i, \alpha - \beta)}(1) = \sum_{\beta \in Q} \sum_{\alpha \in F_q} y_{(i, \alpha - \beta)}(1) = |Q| = k.$$

By a similar analysis, $\sum_{\alpha \in F_q} \vec{u}_{(i, \alpha)}(1) = k \vec{v}_0$ and apply Lemma 2.1 to prove (6) is PSD.

Recall that $p(\Phi)$'s value is $\sum_{j \in [m]} \sum_{f \in C + \vec{b}_j} y_{(C_j, f)}(\vec{1}) = \gamma$, so Φ 's objective value in the w -level Lasserre hierarchy is

$$\begin{aligned} \sum_{j \in [m]} \sum_{f \in C + \vec{b}_j} z_{(C_j, f)}(\vec{1}) &= \sum_{j \in [m]} \sum_{f \in C + \vec{b}_j} \sum_{f' \in F_q^d : f \in f' \oplus Q} y_{(C_j, f')}(\vec{1}) \\ &= \sum_{j \in [m]} \sum_{f' \in F_q^d} \sum_{f \in C + \vec{b}_j} y_{(C_j, f')}(\vec{1}) \cdot 1_{f \in f' \oplus Q} \\ &\geq \sum_{j \in [m]} \sum_{f' \in C + \vec{b}_j} y_{(C_j, f')}(\vec{1}) \cdot |(f' \oplus Q) \cap (C + \vec{b}_j)| \\ &\geq \sum_{j \in [m]} \sum_{f' \in C + \vec{b}_j} y_{(C_j, f')}(\vec{1}) \cdot p_0 |C| \\ &\geq p_0 |C| \cdot \gamma. \end{aligned}$$

□

Before proving Lemma 3.1, We restate Theorem G.8 that is summarized by Chan in [Cha13] of the previous works [Gri01, Sch08, Tul09] and observe that the pseudo-distribution in their construction is uniform over C on every constraint.

Theorem 3.6 ([Cha13]) *Let F_q be the finite field with size q and C be a pairwise independent subspace of F_q^d for some constant $d \geq 3$. The CSP is specified by parameters $F_q, d, k = 1$ and a predicate C . The value of an instance Φ of this CSP on n variables with m constraints is m in the $\Omega(t)$ -level Lasserre hierarchy if every subset T of at most t constraints contains at least $(d - 1.4)|T|$ variables.*

Observation 3.7 *Let $y_S(\{0, 1\}^S)$ denote the pseudo-distribution on S provided by the solution of the semidefinite programming in the Lasserre hierarchy of Φ . For every constraint $C_j (j \in [m])$ in Φ , $y_{C_j}(\{0, 1\}^{C_j})$ provides a uniform distribution over all assignments that satisfy constraint C_j .*

Proof of Lemma 3.1: Without lose of generality, we assume $[N] = [n] \times F_q$. It is natural to think $[N]$ corresponding to n variables and each variables has q vertices corresponding to F_q . Let G be a random

bipartite graph on $[N] \cup [M]$ that is d -regular on $[M]$. For each vertex $j \in M$, the probability that j has two or more neighbors in $i \times F_q$ for some i is at most $\frac{d^2 q}{n}$. Let R denote the subset in M that do not have two or more neighbors in any $i \times F_q$ for all $i \in [n]$. With probability at least $1 - \frac{1}{\sqrt{n}}$, $R \geq (1 - \frac{d^2 q}{\sqrt{n}})M$.

Because the neighbors of each vertex in $[M]$ is generated by choosing d random vertices in $[N]$. For each vertex in R , the generation of its neighbors is as same as first sampling d random variables in $[n]$ then sampling an element in F_q for each variable. By a standard calculation using Chernoff bound and Stirling formula, there exists a constant $\beta = O_{d,M/n}(1)$ such that with high probability, $\forall T \subseteq \binom{R}{\leq \beta n}$, T contains at least $(d - 1.4)|T|$ variables.

We construct an instance Φ based on the induced graph of $[n] \times F_q \cup R$ in the list CSP with the parameters k, q, d and the predicate $\{\vec{0}\}$. For each vertex $j \in R$, let $(i_1, b_1), \dots, (i_d, b_d)$ be its neighbors in G . We add a constraint C_j in Φ with variables x_{i_1}, \dots, x_{i_d} and $\vec{b} = (b_1, \dots, b_d)$.

Recall that C is a subspace staying a subset Q of size k with probability p_0 , we use the following two claims to prove the value of the vertex expansion of ρN -subsets in the Lasserre hierarchy is at most $(1 - p_0)R + (M - R) \leq (1 - p_0)(1 - \frac{d^2 q}{\sqrt{n}})M + \frac{d^2 q}{\sqrt{n}}M \leq (1 - p_0 + o(1))M$ with high probability.

Claim 3.8 Φ 's value is at least $p_0|R|$ in the $\Omega(\beta n)$ -level Lasserre hierarchy.

Claim 3.9 Suppose Φ 's value is at least r in the t -level Lasserre hierarchy, the objective value of the t -level Lasserre hierarchy is at most $|R| - r$ for the vertex expansion problem on $[N] \cup R$ with $\rho = 1 - k/q$.

□

Proof of Claim 3.8: Let Λ be the list CSP with parameters F_q, k, d and predicate C . Let Φ' be the instance of Φ in Λ . From Theorem 3.6, $P(\Phi')$'s value is R because every small constraint subset contains at least $(d - 1.4)|T|$ variables. From Lemma 3.5, Φ 's value is at least $p_0|C| \cdot R$ in the $\Omega(n)$ -level Lasserre hierarchy.

Let us take a closer look, for each constraint j in $P(\Phi')$, the pseudo-distribution on C_j is uniformly distributed over $b_j + C$. Therefore every assignment $f + b_j$ for $f \in C$ appears in the pseudo-distribution of $P(\Phi')$ on C_j with probability $1/|C|$. As the same reason, every assignment $f + b_j$ appears in the pseudo-distribution of Φ' with the same probability $\frac{|Q^d \cap C|}{|C|} = p_0$. Because $\vec{0} \in C$, the probability C_j contains $\vec{0} + \vec{b}_j$ in the pseudo-distribution of Φ' is p_0 by the analysis. Using the solution of Φ' in the $\Omega(\beta n)$ -level Lasserre hierarchy as the solution of Φ , it is easy to see Φ 's value is at least $p_0|R|$. □

Proof of Claim 3.9: Let $y_S(f), \vec{v}_S(f)$ for all $S \subseteq \binom{[n] \times F_q}{t}$ and $f \in \{0, 1\}^S$ be the solution of pseudodistribution and vectors in the t -level Lasserre hierarchy for Φ . We define $z_S(f), \vec{u}_S(f)$ for all $S \subseteq \binom{[n] \times F_q}{t} ([N] = [n] \times F_q)$ and $f \in \{0, 1\}^S$ to be the pseudodistribution and vectors for the vertex expansion problem as follows:

$$\vec{u}_S(f) = \vec{v}_S(\vec{1} - f), z_S(f) = y_S(\vec{1} - f).$$

The verification of the fact that \vec{u} explains the matrix (3) of z in the Lasserre hierarchy is straightforward. Another property from the construction is

$$\sum_{(x_i, b)} \vec{u}_{(x_i, b)}(1) = \sum_{(x_i, b)} \vec{v}_{(x_i, b)}(0) = \sum_{(x_i, b)} (\vec{v}_{\emptyset} - \vec{v}_{(x_i, b)}(1)) = \sum_{i \in [n]} \sum_{b \in F_q} (\vec{v}_{\emptyset} - \vec{v}_{(x_i, b)}(1)) = \sum_i (q\vec{v}_{\emptyset} - k\vec{v}_{\emptyset}) = \rho N \cdot \vec{v}_{\emptyset},$$

which implies the matrix in (4) is positive semidefinite by Lemma 2.1.

The value of the vertex expansion problem given z, \vec{u} is $\sum_{j \in [R]} (1 - z_{N(j)})(\vec{0}) = \sum_{j \in [R]} (1 - y_{(N(j))})(\vec{1}) = R - \sum_{j \in [R]} y_{(N(j))}(\vec{1}) = R - r$. \square

On the other hand, it is easy to prove a random bipartite graph has very good vertex expansion by using Chernoff bound and union bound.

Lemma 3.10 *For any constants $d, \rho, \epsilon > 0$, and $c \geq \frac{20q}{(1-\rho)^d \cdot \epsilon^2}$, with high probability, a random bipartite graph on $[N] \cup [M]$ ($M = cN$) that is d -regular in $[M]$ guarantees that every ρN -subset in $[N]$ contains at least $1 - (1 + \epsilon)(1 - \rho)^d$ different vertices in $[M]$.*

Proof. For any subset $S \subseteq [N]$ of size ρN , the probability that a vertex in $[M]$ is not a neighbor of S is at most $(1 - \rho)^d + o(1)$. Applying Chernoff bound on M independent experiments, the probability that S contains less than $(1 - (1 + \epsilon)(1 - \rho)^d)$ neighbors in $[M]$ is at most $\exp(-\epsilon^2(1 - \rho)^d M / 12) \leq 2^{-M}$. From union bound, every ρN subset has at least $(1 - (1 + \epsilon)(1 - \rho)^d)$ neighbors with high probability. \square

3.1 Integrality gaps for the disperser problem

Theorem 3.11 *For any $\epsilon > 0$ and $\rho \in (0, 1)$, there exist infinitely many d such that a random bipartite graph on $[N] \cup [M]$ that is d -regular in $[M]$ satisfies the following two properties with high probability:*

1. *It is a $(\rho N, (1 - (1 - \rho)^d - \epsilon)M)$ -disperser.*
2. *The objective value of the $\Omega(N)$ -level Lasserre hierarchy for ρ is at most $(1 - C_0 \cdot \frac{1-\rho}{d\rho+1-\rho})M$ for a universal constant $C_0 \geq 1/10$.*

Proof. Let $M \geq \frac{20q}{(1-\rho)^d \cdot \epsilon^2} N$, a random bipartite graph G is a $(\rho N, (1 - (1 - \rho)^d - \epsilon)M)$ -disperser from Lemma 3.10 with very high probability.

On the other hand, choose a prime power q and k in the base cases of Lemma 2.3 or Lemma 2.4 such that $\rho' = 1 - k/q > \rho$ and p_0 be the probability that the subspace C staying in a k -subset. From the construction, $p_0 \geq \frac{1}{3} \frac{1-\rho'}{d\rho'+1-\rho'} \geq \frac{1}{9} \cdot \frac{1-\rho}{d\rho+1-\rho}$. From Lemma 3.1, a random graph G that is d -regular in $[M]$ has vertex expansion at most $(1 - p_0)M$ for ρ' with high probability. Because $\rho' \geq \rho$, this indicates The objective value of the $\Omega(N)$ -level Lasserre hierarchy for ρ is at most $(1 - \frac{1}{9} \cdot \frac{1-\rho}{d\rho+1-\rho})M$. Therefore, a random bipartite graph G satisfies the two properties with high probability. \square

We generalize the above construction to $d = \Theta(\log N)$ and prove the Lasserre hierarchy cannot approximate the entropy of a disperser in the rest of this section. Because $d = \Theta(\log N)$ is a super constant, we relax the strong requirement in the variable expansion of constraints and follow the approach of [Tul09]. We also notice the same observation has independently provided by Bhaskara et.al. in [BCV⁺12].

Theorem 3.12 (Restatement of Theorem 4.3 in [Tul09]) *Let C be the dual space of a linear codes with dimension d and distance l over F_q . Let Φ with n variables and m constraints be an instance of the CSP Λ with $d, k = 1, F_q$ and predicate C . If for every subset S of at most t constraints in Φ , it contains at least $(1 - l/2 + .2)d \cdot |S|$ different variables. Then the value of Φ is m in the $\Omega(t)$ -level Lasserre hierarchy.*

Lemma 3.13 *For any prime power $q, \epsilon > 0, \delta > 0$, and any constant c , a random bipartite graph on $[N] \cup [M]$ that is $d = c \log N$ -regular in M has the following two properties with high probability:*

1. *It is a $(\delta N, (1 - 2(1 - \delta)^d)M)$ -disperser.*

2. The objective value of the $N^{\Omega(1)}$ -level Lasserre hierarchy for $\rho = \frac{q-1}{q}$ is at most $(1 - q^{-\epsilon d} + \frac{1}{N^{1/3}})M$.

Proof. Let A be a linear code over F_q with dimension d , rate $(1 - \epsilon)d$ and distance $3\gamma d$ for some $\gamma > 0$. C is the dual space of A with size $|C| = q^{\epsilon d}$. Let $M = \frac{20q \cdot N}{(1-\delta)^d}$, which is $\text{poly}(N)$ here. From Lemma 3.10, a random bipartite graph G on $[N] \cup [M]$ that is d -regular in M is a $(\delta N, (1 - 2(1 - \delta)^d)M)$ -disperser with very high probability.

In the rest of proof, it is enough to show that for every subsets $S \subseteq \binom{[M]}{\leq N^{\gamma/2}}$ in Φ , the constraints in S contain at least $(1 - \gamma)|S|d$ variables. By union bound, the probability that does not happen is bounded by

$$\sum_{l=1}^{N^{\gamma/2}} \binom{M}{l} \binom{N}{(1-\gamma)d \cdot l} \left(\frac{(1-\gamma)dl}{N} \right)^{dl} \leq \sum_{l \leq N^{\gamma/2}} M^l N^{(1-\gamma)dl} \left(\frac{dl}{N} \right)^{dl} \leq \sum_{l \leq N^{\gamma/2}} \frac{M^l}{N^{\gamma \cdot dl/2}} \frac{(dl)^{dl}}{N^{\gamma \cdot dl/2}} \leq 0.1.$$

By Lemma 3.1, the value of G with $\rho = \frac{q-1}{q}$ is at most $(1 - 1/|C| + \frac{d^2 q}{\sqrt{N}})M \leq (1 - q^{-\epsilon d} + \frac{1}{N^{1/3}})M$ in the $\Omega(n^{\gamma/2})$ -level Lasserre hierarchy. \square

We show the equivalence between the vertex expansion problem and the problem of approximating the entropy in a disperser:

Problem 3.14 Given a bipartite graph $([N], [M], E)$ and ρ , determine the size of the smallest neighbor set over all subsets of size at least ρN in $[N]$.

Problem 3.15 Given a bipartite graph $([N], [M], E)$ and γ , determine the size of the largest subset in $[N]$ with a neighbor set of size $\leq \gamma M$.

We prove the equivalence of these two problems with parameters $\rho + \gamma = 1$. For a bipartite graph $([N], [M], E)$ and a parameter γ , let T be the largest subset in $[N]$ with $|\Gamma(T)| \leq \gamma M$. Let $S = [M] \setminus \Gamma(T)$. Then $|S| \geq (1 - \gamma)M$ and $\Gamma(S) \subseteq [N] \setminus T$. Since T is the largest subset with $|\Gamma(T)| \leq \gamma M$, S is the subset of size at least $(1 - \gamma)M$ with the smallest neighbor set. The converse is similar, which shows the equivalence between these two problems.

Theorem 3.16 For any $\alpha \in (0, 1)$, any $\delta \in (0, 1)$ and any prime power q , there exists a constant c such that a random bipartite graph on $[N] \cup [M]$ that is $D = c \log N$ -regular in $[N]$ has the following two properties with high probability:

1. It is an $(N^\alpha, (1 - \delta)M)$ -disperser.
2. The objective value of the SDP in the $N^{\Omega(1)}$ -level Lasserre hierarchy for obtaining M/q distinct neighbors is at least $N^{1-\alpha/2}$.

Proof. Let $\epsilon = \frac{\log \frac{1}{1-\delta}}{4\alpha \log q} = O(1)$ and $d = \frac{\log N}{4\epsilon \log q}$ such that $|C| = q^{\epsilon d} = N^{1/4}$ and $M = \frac{20q \cdot N}{(1-\delta)^d} \geq N^{1/\alpha}$. So $d = O(\log M)$.

From Lemma 3.13, a random bipartite graph on $[N] \cup [M]$ d -regular in $[M]$ is a $(\delta N, M - M^\alpha)$ -disperser, but the value of $N^{\Omega(1)}$ -level Lasserre hierarchy for G with $\rho = 1 - 1/q$ is at most $M - M^{1-\alpha/2}$. From the equivalence, any subset of size M^α in $[M]$ has a neighbor set of size at least $(1 - \delta)N$. On the other hand, it is possible that there exists a $M^{1-\alpha/2}$ -subset of $[M]$ with a neighbor set of size at most $[N]/q$ in the Lasserre hierarchy, from the fact that the $N^{\Omega(1)}$ -level Lasserre hierarchy has a value at most $M - M^{1-\alpha/2}$ for $\rho = 1 - 1/q$. To finish the proof, swap $[N]$ and $[M]$ in the bipartite graph such that $D = d$ in the new bipartite graph. \square

Corollary 3.17 (Restatement of Theorem 1.3) *For any $\alpha \in (0, 1)$, any $\delta \in (0, 1)$, there exists a constant c such that a random bipartite graph on $[N] \cup [M]$ with $D = c \log N$ -regular in $[N]$ has the following two properties with high probability:*

1. *It is an $(N^\alpha, (1 - \delta)M)$ -disperser.*
2. *The objective value of the SDP in the $N^{\Omega(1)}$ -level Lasserre hierarchy for obtaining δM distinct neighbors is at least $N^{1-\alpha}$.*

3.2 An integrality gap for the expander problem

We prove that a random bipartite graph is almost D -regular on the right hand side and use the fact $dN \approx DM$.

Theorem 3.18 *For any prime power q , integer $d < q$ and constant $\delta > 0$, there exist a constant D and a bipartite graph G on $[N] \cup [M]$ with the largest left degree D and the largest right degree d has the following two properties for $\rho = 1/q$:*

1. *It is a $(\rho N, (1 - \epsilon' - 2\delta)D)$ -expander with $\epsilon' = \frac{(1-\rho)^d - (1-\rho d)}{\rho d} = \sum_{i=1}^{d-1} (-1)^{i-1} \frac{(d-1) \cdots (d-i+1)}{(i+1)!} \rho^i$.*
2. *The objective value of the vertex expansion for G with ρ in the $\Omega(N)$ -level Lasserre hierarchy is at most $(1 - \epsilon + \delta)D \cdot \rho N$ with $\epsilon = \frac{\rho(d-1)}{2}$.*

Proof. Let β be a very small constant specified later and $c = \frac{100q \cdot \log(1/\beta)}{d(1-\rho)^d \cdot \delta^2}$. Let G_0 be a random graph on $[N] \cup [M]$ with $M = cN$ that is d -regular in $[M]$. Let $D_0 = \frac{dM}{N}$ and L denote the vertices in $[N]$ with degree $[(1 - \delta)D_0, (1 + \delta)D_0]$. Let G_1 denote the induced graph of G_0 on $L \cup [M]$. The largest degree of L is $D = (1 + \delta)D_0$ and the largest degree of M is d . We will prove G_1 is a bipartite graph that satisfies the two properties in this lemma with high probability. Because G_0 is a random graph, we assume there exists a constant $\gamma = O_{M/N, d}(1)$ such that every subset $S \in \binom{M}{\leq \gamma N}$ has different $(d - 1.1)|S|$ neighbors.

In expectation, each vertex in N has degree D_0 . By Chernoff bound, the fraction of vertices in $[N]$ of G_0 with degree more than $(1 + \delta)D_0$ or less than $(1 - \delta)D_0$ is at most $2 \exp(-\delta^2 \cdot \frac{d}{N} \cdot M/12) \leq \beta^4$. At the same time, with high probability, G_0 satisfies that any $\beta^3 N$ -subset in $[N]$ has total degree at most βdM because $\binom{N}{\beta^3 N} \cdot \exp(-(\frac{1}{\beta^2})^2 (\beta^3 d) \cdot M/12)$ is exponentially small in N . Therefore with high probability, $|L| \geq (1 - \beta^3)N$ and there are at least $(1 - \beta)dM$ edges in G_1 .

We first verify the objective value of the vertex expansion for G_1 with $\rho = 1/q$ in the $\Omega(N)$ -level Lasserre hierarchy is at most $(1 - \epsilon + \delta)D \cdot \rho N$. Let R be the vertices in $[M]$ that have degree d . From Lemma 3.1, the objective value of the vertex expansion for $L \cup R$ with $\rho = 1/q$ in the $\Omega(\gamma N)$ -level Lasserre hierarchy is at most $(1 - p_0)|R|$ where p_0 is the staying probability of C in a $q - 1$ subset. From Lemma 2.4, $p_0 = 1 - d\rho + \binom{d}{2}\rho^2$. Therefore $(1 - p_0)|R| \geq (1 - 1 + d\rho - \binom{d}{2}\rho^2)(1 - d\beta)M$. For the vertices in $M \setminus R$, they will contribute at most $d\beta M$ in the objective value of the Lasserre hierarchy. Therefore the objective value for G_1 is at most $(d\rho - \binom{d}{2}\rho^2 + d\beta)M = (1 - \frac{(d-1)\rho}{2} + \frac{\beta}{\rho})\rho dM \leq (1 - \epsilon + \frac{\beta}{\rho})\rho dM$.

For the integral value, every ρN -subset in $[N]$ has at least $(1 - (1 + \beta)(1 - \rho)^d)M$ neighbors in G_0 by Lemma 3.10. Because G_1 is the induced graph of G_0 on $L \cup [M]$, every ρN -subset in L has at least $(1 - (1 + \beta)(1 - \rho)^d)M \geq (1 - \epsilon' - \frac{\beta}{\rho d})\rho dM \geq (1 - \epsilon' - \frac{\beta}{\rho d})D_0 \cdot \rho N$ neighbors in G_1 . By setting β small enough, there exists a bipartite graph with the required two properties. \square

Corollary 3.19 *For any $\epsilon > 0$ and any $\epsilon' < \frac{e^{-2\epsilon} - (1-2\epsilon)}{2\epsilon}$, there exist ρ small enough and a bipartite graph G with the largest left degree $D = O(1)$ that has the following two properties:*

1. It is a $(\rho N, (1 - \epsilon')D)$ -expander.
2. The objective value of the vertex expansion for G with ρ in the $\Omega(N)$ -level Lasserre hierarchy is at most $(1 - \epsilon)D \cdot \rho N$.

Proof. Think ρ to be a small constant and $d = \frac{2\epsilon}{\rho} + 1$ such that ϵ is very close to $\frac{\rho d}{2}$. Then the limit of $\epsilon' = \frac{(1-\rho)^d - (1-\rho d)}{\rho d}$ is $\frac{e^{-\rho d} - (1-\rho d)}{\rho d} = \frac{e^{-2\epsilon} - (1-2\epsilon)}{2\epsilon}$ by decreasing ρ . \square

4 Approximation Algorithm

In this section, we will provide a polynomial time algorithm that has an approximation ratio close to the integrality gap in Theorem 1.4.

Theorem 4.1 *Given a bipartite graph $([N], [M], E)$ with right degree d , if $(1 - \Delta)M$ is the size of the smallest neighbor set over ρN -subsets in $[N]$, there exists a polynomial time algorithm that outputs a subset $T \subseteq [N]$, such that $|T| = \rho N$ and $\Gamma(T) \leq (1 - \Omega(\frac{\min\{(\frac{\rho}{1-\rho})^2, 1\}}{\log d}) \cdot d(1 - \rho)^d \cdot \Delta)M$.*

We consider a simple semidefinite programming for finding a subset $T \subseteq [N]$ that maximizes the number of unconnected vertices to T .

$$\begin{aligned} & \max \sum_{j \in [M]} \left\| \frac{1}{d} \sum_{i \in \Gamma(j)} \vec{v}_i \right\|_2^2 \\ \text{Subject to } & \langle \vec{v}_i, \vec{v}_i \rangle \leq 1 \\ & \sum_{i=1}^n \vec{v}_i = \vec{0} \end{aligned} \quad (*)$$

We first show the objective value of the semidefinite programming is at least $\min\{(\frac{\rho}{1-\rho})^2, 1\} \cdot \Delta$. For convenience, let δ denote the value of this semidefinite programming and A denote the positive definite matrix of the objective function in the semidefinite programming such that $\delta = \sum_{i,j} A_{i,j}(\vec{v}_i^T \cdot \vec{v}_j)$. If $\rho \geq 0.5$, $\delta \geq \Delta \cdot M$ by choosing $\vec{v}_i = (1, 0, \dots, 0)$ for every $i \notin S$ and $\vec{v}_i = (-\frac{1-\rho}{\rho}, 0, \dots, 0)$ for every $i \in S$. But this is not a valid solution for the SDP when $\rho < 0.5$. However, $\delta \geq (\frac{\rho}{1-\rho})^2 \cdot \Delta \cdot M$ in this case by choosing $\vec{v}_i = (\frac{\rho}{1-\rho}, 0, \dots, 0)$ for every $i \notin S$ and $\vec{v}_i = (-1, 0, \dots, 0)$ for every $i \in S$. Therefore $\delta \geq \min\{(\frac{\rho}{1-\rho})^2, 1\} \cdot \Delta M$. Without loss of generality, both δ and ΔM are $\geq \frac{1}{d}M$, otherwise a random subset is enough to achieve the desired approximation ratio.

The algorithm has two stages: first round \vec{v}_i to $z_i \in [-1, 1]$ and keep $\sum_i z_i$ almost balanced, which is motivated by the work [AN04], then round z_i to x_i using the algorithm suggested by [CMM07].

Lemma 4.2 *There exists a polynomial time algorithm that given $\|\vec{v}_i\| \leq 1$ for every i , $\sum_i \vec{v}_i = \vec{0}$ and $\delta = \sum_j \left\| \frac{1}{d} \sum_{i \in \Gamma(j)} \vec{v}_i \right\|_2^2 \geq M/d$, it finds $z_i \in [-1, 1]$ for every i such that $|\sum_i z_i| = O(N/d)$ and $\sum_j (\frac{1}{d} \sum_{i \in \Gamma(j)} z_i)^2 \geq \Omega(\frac{\delta}{\log d})$.*

Proof. The algorithm works as follows:

1. Sample $\vec{g} \sim N(0, 1)^N$ and choose $t = 3\sqrt{\log d}$.
2. Let $\zeta_i = \langle \vec{g}, \vec{v}_i \rangle$ for every $i = 1, 2, \dots, n$.

3. If $\zeta_i > t$ or $\zeta_i < -t$, cut $\zeta_i = \pm t$ respectively.

4. $z_i = \zeta_i/t$.

It is convenient to analyze the approximation ratio in another set of vectors $\{\vec{u}_i | i \in [n]\}$ in a Hilbert space such that $\vec{u}_i(\vec{g}) = \langle \vec{v}_i, \vec{g} \rangle$ and $\langle \vec{u}_i, \vec{u}_j \rangle = E_{\vec{g}}[\langle \vec{u}_i, \vec{g} \rangle \cdot \langle \vec{g}, \vec{u}_j \rangle] = \langle \vec{v}_i, \vec{v}_j \rangle$. So $\sum_{i,j} A_{i,j}(\vec{u}_i^T \cdot \vec{u}_j) = \delta$ and $\sum_i \vec{u}_i = \vec{0}$ again. Let \vec{u}'_i be the vector in the same Hilbert space by applying the cut operation with parameters t on \vec{u}_i . Namely $\vec{u}'_i(\vec{g}) = t$ (or $-t$) when $\vec{u}_i(\vec{g}) > t$ (or $< -t$), otherwise $\vec{u}'_i(\vec{g}) = \vec{u}_i(\vec{g}) \in [-t, t]$. Therefore the algorithm is as same as sampling a random point \vec{g} and setting $z_i = \vec{u}'_i(\vec{g})/t$.

Fact 4.3 For every i , $\|\vec{u}'_i - \vec{u}_i\|_1 = O(1/d^{4.5})$ and $\|\vec{u}'_i - \vec{u}_i\|_2^2 = O(1/d^4)$.

The analysis uses the second fact to bound $\sum_{i,j} A_{i,j}((\vec{u}_i - \vec{u}'_i)^T \cdot \vec{u}_j) \leq O(m/d^2)$ as follows. Notice that A is a positive definite matrix and consider $\sum_{i,j} A_{i,j}(\vec{w}_i^T \cdot \vec{w}_j)$ for any unit vectors \vec{w}_i and \vec{w}_j . It reaches the maximal value when $\vec{w}_i = \vec{w}_j$ by the property of positive definite matrices. And $\sum_{i,j} A_{i,j}(\vec{w}_i^T \cdot \vec{w}_j) = \sum_{j \in [M]} \|\frac{1}{d} \sum_{i \in \Gamma(j)} \vec{w}_i\|_2^2$ is always bounded by M , because \vec{w}_i are unit vectors. So $\sum_{i,j} A_{i,j}(\vec{w}_i^T \cdot \vec{w}_j) \leq \max\{\|\vec{w}_1\|_2, \dots, \|\vec{w}_n\|_2\} \cdot \max\{\|\vec{w}'_1\|_2, \dots, \|\vec{w}'_n\|_2\} \cdot M$.

$$\begin{aligned} & \sum_{i,j} A_{i,j}(\vec{u}_i^T \cdot \vec{u}_j) - \sum_{i,j} A_{i,j}(\vec{u}'_i^T \cdot \vec{u}_j) \\ &= \sum_{i,j} A_{i,j}(\vec{u}_i^T \cdot (\vec{u}_j - \vec{u}'_j)) + \sum_{i,j} A_{i,j}((\vec{u}_i - \vec{u}'_i)^T \cdot \vec{u}_j) \\ &\leq O(M/d^2) \end{aligned}$$

Therefore $\sum_{i,j} A_{i,j}(\vec{u}'_i^T \cdot \vec{u}_j) \geq 0.99 \sum_{i,j} A_{i,j}(\vec{u}_i^T \cdot \vec{u}_j) \geq 0.99M/d$. And it is upper bounded by $t^2 \cdot M$. So with probability at least $\frac{49}{dt^2}$, g satisfies $\sum_{i,j} A_{i,j}(\vec{u}'_i(g) \cdot \vec{u}_j(g)) \geq .49\delta$. On the other hand, $|\sum_i \vec{u}'_i(g)| \geq N/d$ with probability at most $1/d^3$ from the first property $\|\sum_i \vec{u}'_i\|_1 \leq O(N/d^4)$.

Overall, with probability at least $\frac{5}{dt^2} - 1/d^3$, z_i satisfies $|\sum z_i| = O(N/d)$ and $\sum_j (\frac{1}{d} \sum_{i \in \Gamma(j)} z_i)^2 = \Omega(\frac{\delta}{t^2}) = \Omega(\frac{\delta}{\log d})$. \square

It is not difficult to verify that independently sampling $z_i \in \{-1, 1\}$ for every i according to its bias z_i will not reduce the objective value but keep the same bias overall i . Without lose of generality, let $z_i \in \{-1, 1\}$ from now on.

Lemma 4.4 There exists a polynomial time algorithm that given z_i with $|\sum_i z_i| = O(N/d)$, outputs $x_i \in \{0, 1\}$ such that $\sum_i x_i = (1 - \rho)(1 \pm 1/d^{1.5})N$ and $\sum_j 1_{\forall i \in \Gamma(j): x_i=1} \geq \Omega(d(1 - \rho)^d) \cdot \sum_j (\frac{1}{d} \sum_{i \in \Gamma(j)} z_i)^2$.

Proof. The algorithm works as follows:

1. $\delta = (1 - \rho)\sqrt{2/d}$. Execute Step 2 or Step 3 with probability 0.5 and 0.5 separately.
2. For every $i \in [N]$, $x_i = 1$ with probability $1 - \rho + \delta z_i$.
3. For every $i \in [N]$, $x_i = 1$ with probability $1 - \rho - \delta z_i$.

Let $y_j = \frac{1}{d} \sum_{i \in \Gamma(j)} z_i$. The probability $x_i = 1$ for every i in $\Gamma(j)$ is

$$\begin{aligned}
& \frac{1}{2} \left((1 - \rho + \delta)^{\frac{1+y_j}{2}d} \cdot (1 - \rho - \delta)^{\frac{1-y_j}{2}d} + (1 - \rho - \delta)^{\frac{1+y_j}{2}d} \cdot (1 - \rho + \delta)^{\frac{1-y_j}{2}d} \right) \\
&= \frac{1}{2} (1 - \rho + \delta)^{d/2} (1 - \rho - \delta)^{d/2} \left(\left(\frac{1 - \rho + \delta}{1 - \rho - \delta} \right)^{y_j \cdot d/2} + \left(\frac{1 - \rho - \delta}{1 - \rho + \delta} \right)^{y_j \cdot d/2} \right) \\
&= \frac{1}{2} (1 - \rho)^d \left(1 - \frac{\delta^2}{(1 - \rho)^2} \right)^{d/2} \cdot \cosh(y_j \cdot d/2 \cdot \ln(\frac{1 - \rho + \delta}{1 - \rho - \delta})) \\
&\geq \frac{1}{2} (1 - \rho)^d (1 - 2/d)^{d/2} \cdot 0.9 \cdot (y_j \cdot d/2 \cdot \ln(\frac{1 - \rho + (1 - \rho)\sqrt{2/d}}{1 - \rho - (1 - \rho)\sqrt{2/d}}))^2 \\
&\geq \frac{1}{2} (1 - \rho)^d (1 - 2/d)^{d/2} \cdot 0.9 \cdot y_j^2 \cdot (d/2)^2 \cdot (\sqrt{2/d})^2 \\
&\geq \Omega((1 - \rho)^d \cdot y_j^2 \cdot d).
\end{aligned}$$

At the same time, $\sum_i x_i$ is concentrated around $E[\sum_i x_i] = \sum_i (1 - \rho \pm \delta z_i) = (1 - \rho)N \pm \delta \sum_i z_i = (1 - \rho)(1 \pm 1/d^{1.5})N$ with very high probability. Therefore $\{x_1, \dots, x_n\}$ satisfies $\sum_i x_i = (1 - \rho)(1 \pm 1/d^{1.5})N$ and $\sum_j 1_{\forall i \in \Gamma(j): x_i = 1} \geq \Omega(d(1 - \rho)^d) \cdot \sum_j y_j^2$ with constant probability. \square

Proof. [Proof of Theorem 4.1] Let δ be the value from SDP (*), which is $\geq \min\{(\frac{\rho}{1-\rho})^2, 1\} \cdot \Delta$ from the analysis above. By Lemma 4.2, round v_i into $z_i \in [-1, 1]$ such that $|\sum_i z_i| = O(N/d)$ and $\sum_j (\frac{1}{d} \sum_{i \in \Gamma(j)} z_i)^2 \geq \Omega(\frac{\delta}{\log d})$. By Lemma 4.4, round z_i into $x_i \in \{0, 1\}$ such that $\sum_i x_i = (1 - \rho)(1 \pm 1/k^{1.5})N$ and $\sum_j 1_{\forall i \in \Gamma(j): x_i = 1} \geq \Omega(d(1 - \rho)^d \cdot \frac{\delta}{\log d})$.

Let $T = \{i | x_i = 0\}$. Then $|T| = \rho(1 \pm O(\frac{1}{k^{1.5}}))N$ and $\Gamma(T) \leq (1 - C \cdot \frac{\min\{(\frac{\rho}{1-\rho})^2, 1\}}{\log d}) \cdot d(1 - \rho)^d \cdot \Delta)M$ for some absolute constant C . At last, adjust the size of T by randomly adding or deleting $O(\frac{N}{k^{1.5}})$ vertices such that the size of T is ρN . Because at most $O(\frac{N}{k^{1.5}})$ vertices are added to T , with constant probability, $\Gamma(j) \cap T = \emptyset$ if $\Gamma(j) \cap T = \emptyset$ for a node $j \in [M]$ before the adjustment. Therefore $\Gamma(T) \leq (1 - C_0 \cdot \frac{\min\{(\frac{\rho}{1-\rho})^2, 1\}}{\log d}) \cdot d(1 - \rho)^d \cdot \Delta)M$ for some absolute constant C_0 . \square

5 Hardness and Approximation for $(\rho N, \rho(1 + \epsilon)M)$ -disperser

In this section, we assume $G = ([N], [M], E)$ is D -regular on left and d -regular on right. We present our results for $(\rho N, \rho(1 + \epsilon)M)$ -dispersers when ϵ is small enough. We first make a reduction from vertex expansion [LRV13] to the disperser problem which gives a hardness result based on Small-Set Expansion hypothesis. Then we show there is a polynomial time algorithm that has an approximation ratio close to the hardness result when $d|D$. Let e be the base of the natural logarithm in this section.

Theorem 5.1 (Restatement of Theorem 1.3 in [LRV13]) *For every $\eta > 0$ and $\rho = \frac{1}{e^q}$ for a natural number q , there exists an absolute constant C_0 such that $\forall \epsilon > 0$ it is SSE hard to distinguish between the following two cases for a given graph $H = (V, E)$ with maximal degree $d = O(1/\epsilon)$:*

1. *There exists a set $S \subset V$ of size $\rho|V|$ such that $|\Gamma(S) \setminus S| \leq \epsilon \cdot |S|$.*
2. *For every subset $S \subset V$ of size $\leq \frac{1}{2}|V|$, $|\Gamma(S) \setminus S| \geq (\min\{10^{-10}, C_0\sqrt{\epsilon \log d}\} - \eta)|S|$.*

Remark 5.2 In [LRV13], Louis et.al. proved there exists a subset of size $\rho = \frac{1}{2e}$ in the complete case. Their construction can be generalized to $\rho = \frac{1}{q \cdot e}$ by enlarging the alphabet from $\{0, 1\}$ to $[q]$.

Theorem 5.3 For every small δ and $C > 1$, there exist a small constant γ and a large integer D such that it is SSE hard to distinguish a bipartite graph on $[N] \cup [M]$ with left degree D is between the following two cases:

1. There exists a set $S \subset V$ of size γN such that $|\Gamma(S)| \leq (1 + \delta) \cdot |S|$.
2. For every subset $S \subset V$ of size γN , $|\Gamma(S)| \geq C|S|$.

Proof. Let $\epsilon = \delta^3$, $\rho = \frac{1}{q \cdot e} < \frac{\delta}{4c}$ for some integer q and $k = \frac{1}{2\delta^2}$. We start from a graph H in Theorem 5.1 to amplify the gap between $1 + \epsilon$ and $1 + \sqrt{\epsilon \log d}$. Let $|V| = n$ and $V_1 = V_2 = V$ in the bipartite graph $G_0 = (V_1, V_2, E)$. There is an edge $(i, j) \in E$ between $i \in V_1$ and $j \in V_2$ iff (i, j) is also an edge in G or $i = j$. For any subset $S \subseteq V_1$, $|\Gamma(S)| = |\Gamma_G(S) \setminus S| + |S|$ because $\Gamma(S) = \Gamma_H(S) \cup S$ in the construction of G_0 .

Let k Let $G_1 = (V_1^k, V_2^k \cup (V_1 \times [W]), E')$ where $W = \frac{\delta}{2} \rho^{k-1} n^{k-1}$. There is a edge between $(a_1, a_2, \dots, a_k) \in V_1^k$ and $(b_1, \dots, b_k) \in V_2^k$ if and only if for every $i \in [k]$, $(a_i, b_i) \in E$ of G_0 or $a_i = b_i$. There is a edge between $(a_1, a_2, \dots, a_k) \in V_1^k$ and $(i, w) \in V_1 \times [W]$ iff $i \in \{a_1, \dots, a_k\}$.

In the completeness case, there is a subset S of size ρN such that $\Gamma_{G_0}(S) \leq (1 + \epsilon)|S|$. So $\Gamma_{G_1}(S^k) \leq (1 + \epsilon)^k |S|^k + (1 + \epsilon)|S|W = (1 + \delta/2)\rho^k n^k + (1 + \epsilon)\rho n W \leq (1 + \delta)\rho^k n^k$.

For the sound case, let T be an $(\rho n)^k$ -subset of V_1^k . there are two cases: one is that each coordinate expands at least $(1 + \sqrt{\epsilon})$ as the soundness of G_0 . Otherwise it reach $\frac{1}{2}|V|$ such that it stops expanding at some moment. So $|\Gamma(S)| = \min\{(1 + \sqrt{\epsilon})^k |T| + (1 + \sqrt{\epsilon})\rho n W, \frac{n}{2}W\} \geq C|T|$ from our choices of parameters. \square

Our algorithm is based on the approximation algorithm of Louis and Makarychev in [LM14]. We modify their algorithm and outline the analysis. For a graph $H = (V, E)$, let $\phi_H(S) = \frac{|\Gamma_H(S) \setminus S|}{|S|}$ for $S \subset V$ and $\phi_{H,\rho} = \min_{S \subset V: |S| \leq \rho N} \{\phi_H(S)\}$.

Theorem 5.4 (Restatement of Theorem 1.8 in [LM14]) There is a polynomial time algorithm for the Small Set Vertex Expansion problem that for any $\delta > 0$ given a graph $H = (V, E)$ with maximal degree d' , finds a set $S \subset V$ of size at most $(1 + \delta)\rho|V|$ such that $\phi_H(S) \leq O_\delta(\sqrt{\phi_{H,\rho} \cdot d'} \cdot \frac{1}{\rho} \log \frac{1}{\rho} \log \log \frac{1}{\rho} + \epsilon/\rho)$.

It is not difficult to generalize this result with a k -to-1 map ψ from V to W : redefine $\phi_H^\psi(T) = \phi_H(\psi^{-1}(T))$ for $T \subset W$ and $\phi_{H,\rho}^\psi = \min_{T \subset W: |T| \leq \rho|W|} \{\phi_H^\psi(T)\}$.

Corollary 5.5 There is a polynomial time algorithm for the Small Set Vertex Expansion problem that for any $\delta > 0$ given a graph $H = (V, E)$ with maximal degree d' and a k -to-1 map ψ from V to W , finds a set $T \subset W$ of size at most $(1 + \delta)\rho|W|$ such that $\phi_H^\psi(T) \leq O_\delta(\sqrt{\phi_{H,\rho}^\psi \cdot d'} \cdot \frac{1}{\rho} \log \frac{1}{\rho} \log \log \frac{1}{\rho} + \epsilon/\rho)$.

We briefly explain why the algorithm in [LM14] works with ψ . In [LRV13], Louis et.al. prove the equivalence of vertex expansion and symmetric vertex expansion, which is defined to be $\phi^V(S) = \frac{|\Gamma(S) \setminus S| + |\Gamma(\bar{S}) \setminus \bar{S}|}{\min\{|S|, |\bar{S}|\}}$. In [LM14], Louis and Makarychev relax symmetric vertex expansion as a semidefinite program with L_2^2 inequality. The main tool in the algorithm [LM14] is orthogonal separators which is

introduced by [CMM06,BFK⁺11]. We modify the SDP by adding an extra constrain (**):

$$\begin{aligned}
& \min \sum_{j \in V} \max_{j_1 \in \Gamma(j), j_2 \in \Gamma(j)} \{\|\vec{v}_{j_1} - \vec{v}_{j_2}\|_2^2\} \\
& \text{Subject to } \sum_j \|\vec{v}_j\|_2^2 = 1 \\
& \sum_k \langle \vec{v}_j, \vec{v}_k \rangle \leq \rho |V| \cdot \|\vec{v}_j\|_2^2 \quad \forall j \in V \\
& \|\vec{v}_k - \vec{v}_j\|_2^2 + \|\vec{v}_l - \vec{v}_j\|_2^2 \geq \|\vec{v}_k - \vec{v}_l\|_2^2 \quad \forall k, l, j \in V \\
& 0 \leq \langle \vec{v}_j, \vec{v}_k \rangle \leq \|\vec{v}_j\|_2^2 \quad \forall j, k \in V \\
& \|\vec{v}_j - \vec{v}_k\|_2^2 = 0 \quad \forall j, k \in V : \psi(j) = \psi(k) \quad (**)
\end{aligned}$$

We apply the same rounding process in Theorem 5.4 and choose $i \in [W]$ in T or not according to the value of an arbitrary element in $\psi^{-1}(i)$. The orthogonal separator in Theorem 5.4 rounds \vec{v}_j to $x_j \in \{0, 1\}$ for every vector $j \in V$. All vectors in $\psi^{-1}(i)$ for $i \in [W]$ will be rounded into the same value, because $\vec{v}_j = \vec{v}_k$ for all $j, k \in \psi^{-1}(i)$ and the rounding value only depends on the vector in the orthogonal separator. Therefore we choose i to be in T or not according to the value x_j for $j \in \psi^{-1}(i)$. Because ψ is a k -to-1 map, $\frac{|T|}{|W|} = \frac{\sum_j x_j}{|V|}$ and the algorithm has the same guarantees.

Theorem 5.6 *There exists a polynomial time algorithm that given a regular bipartite graph $G = ([N], [M], E)$ with $d|D$ that is not a $(\rho N, \rho(1 + \epsilon)M)$ -disperser, finds a subset S with size $(1 \pm \delta)\rho N$ and $\Gamma(S) \leq (1 + O_\delta(\sqrt{\epsilon \log(d + D/d)} \cdot \frac{1}{\rho} \cdot \log \frac{1}{\rho} \log \log \frac{1}{\rho} + \epsilon \cdot \rho^{-1}))|S|$.*

Proof. Recall D and d are the left and right degree of $G = ([N], [M], E)$ respectively. Let $k = M/N = D/d$. At first, there exists a k -to-1 map ψ from $[M]$ to $[N]$ by Hall's theorem. It is enough to consider an equivalent problem of vertex expansion on $H = ([M], E)$ that is not necessarily bipartite:

1. For every vertex $j \in [N]$, partition $\Gamma(j)$ into k groups N_1, N_2, \dots, N_k .
2. Let $\{i_1, \dots, i_k\} = \psi^{-1}(j)$.
3. For every $l \in [k]$, connect i_l to every vertex of N_l in H (allow self-edge here). So the degree of every vertex in H is $2d$.
4. The new problem in H is to find a subset $T \subset [N]$ with size $\leq \rho N$ that minimizes $\phi_H(\psi^{-1}(T))$.

For any subset T in $[N]$, $\Gamma(T)$ in G equals to $\Gamma_H(\psi^{-1}(T))$ in H because if $j \in [M]$ is a neighbor of a vertex i in $T \subseteq [N]$, j is connected to some node of $\psi^{-1}(i)$ in H from the construction. Another property of the construction is $\psi^{-1}(T) \subseteq \Gamma_H(\psi^{-1}(T))$ for any subset $T \subset [N]$.

Because G is not a $(\rho N, \rho(1 + \epsilon)M)$ disperser, there exists $S_0 \subset [N]$ with size ρN and $\Gamma(S_0) \leq \rho(1 + \epsilon)M$ in the bipartite graph. Therefore $\phi_H(\psi^{-1}(S_0)) \leq \frac{\rho \cdot \epsilon M}{\rho M} \leq \epsilon$. Repeat applying the algorithm in Corollary 5.5 at most ρn times to find a subset as follows:

1. Apply the algorithm on H with ψ and ρ to finding a subset $T_0 \leq \rho(1 + \delta)N$ such that $\phi_H^\psi(T_0) \leq O_\delta(\sqrt{\epsilon \cdot d'} \cdot \frac{1}{\rho} \log \frac{1}{\rho} \log \log \frac{1}{\rho} + \epsilon/\rho) \leq \tilde{O}_{\delta, \rho}(\sqrt{\epsilon \cdot 2d} \cdot \frac{1}{\rho})$. If $|T_0| > \rho(1 - \delta)N$, then return $T = T_0$.

2. Otherwise consider $H' = H \setminus T_0$ and $S_1 = S_0 \setminus T_0$, $\phi(\psi^{-1}(S_1)) \leq \frac{|\Gamma(S_0 \setminus T_0) \setminus S_1|}{|S_1|} \leq \frac{|\Gamma(S_0) \setminus S_0|}{|S_1|} \leq \frac{\epsilon}{\delta}$. Then apply the algorithm in Corollary 5.5 on H' with ψ and $\rho' = \rho - \frac{T_0}{N}$ to finding a subset $T_1 \leq \rho'(1 + \delta)N$. If $|T_0 \cup T_1| \geq \rho(1 - \delta)N$, then return $T_0 \cup T_1$.
3. Otherwise consider $H' \setminus T_1$ and $S_1 \setminus T_1$ again.
4. Eventually the algorithm will find a subset $T = T_0 \cup T_1 \cup \dots \cup T_{\rho N}$ such that $|T| \in [\rho(1 - \delta)N, \rho(1 + \delta)N]$ and $\Gamma(T) \subseteq \Gamma(T_0) + \Gamma(T_1) + \dots + \Gamma(T_{\rho N})$. From the guarantee of each T_i , $\phi(\psi^{-1}(T)) \leq O(\sqrt{2d \frac{\epsilon}{\delta}} \cdot \frac{1}{\rho})$.

□

6 Acknowledgement

We are grateful to David Zuckerman for his introduction to this problem, as well as for many fruitful discussions without which this work would not have been completed. Thanks for the suggestions and reviews from anonymous reviewers especially for the suggestion of the name “list Constraint Satisfaction problems”.

References

- [ABG13] Per Austrin, Siavosh Benabbas, and Konstantinos Georgiou. Better balance by being biased: A 0.8776-approximation for max bisection. In *Proceedings of the 24th Annual ACM-SIAM SODA*, pages 277–294, 2013.
- [AN04] Noga Alon and Assaf Naor. Approximating the cut-norm via grothendieck’s inequality. In *Proceedings of the 36th Annual ACM Symposium on Theory of Computing, Chicago, IL, USA, June 13-16, 2004*, pages 72–80, 2004.
- [ARV09] Sanjeev Arora, Satish Rao, and Umesh Vazirani. Expander flows, geometric embeddings and graph partitioning. *J. ACM*, 56(2):5:1–5:37, April 2009.
- [Bar14] Boaz Barak. Sums of squares upper bounds, lower bounds, and open questions. <http://www.boazbarak.org/sos/>, 2014. Page 39. Accessed: October 28, 2014.
- [BCV⁺12] Aditya Bhaskara, Moses Charikar, Aravindan Vijayaraghavan, Venkatesan Guruswami, and Yuan Zhou. Polynomial integrality gaps for strong sdp relaxations of densest k-subgraph. In *Proceedings of the 23rd Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 388–405. SIAM, 2012.
- [BFK⁺11] Nikhil Bansal, Uriel Feige, Robert Krauthgamer, Konstantin Makarychev, Viswanath Nagarajan, Joseph (Seffi) Naor, and Roy Schwartz. Min-max graph partitioning and small set expansion. In *Proceedings of the 2011 IEEE 52nd Annual Symposium on Foundations of Computer Science*, pages 17–26, Washington, DC, USA, 2011. IEEE Computer Society.
- [BGGP12] Itai Benjamini, Ori Gurel-Gurevich, and Ron Peled. On k-wise independent distributions and boolean functions. <http://arxiv.org/abs/1207.0016>, 2012. Accessed: October 28, 2014.

- [BGH⁺12] Boaz Barak, Parikshit Gopalan, Johan Hastad, Raghu Meka, Prasad Raghavendra, and David Steurer. Making the long code shorter. In *Proceedings of the 2012 IEEE 53rd Annual Symposium on Foundations of Computer Science*, pages 370–379, Washington, DC, USA, 2012. IEEE Computer Society.
- [BKS⁺10] B. Barak, G. Kindler, R. Shaltiel, B. Sudakov, and A. Wigderson. Simulating independence: New constructions of condensers, ramsey graphs, dispersers, and extractors. *J. ACM*, 57(4):20:1–20:52, May 2010.
- [BMRV00] H. Buhrman, P. B. Miltersen, J. Radhakrishnan, and S. Venkatesh. Are bitvectors optimal? In *Proceedings of the 32nd Annual ACM Symposium on Theory of Computing, STOC*, pages 449–458, New York, NY, USA, 2000. ACM.
- [BOT02] Andrej Bogdanov, Kenji Obata, and Luca Trevisan. A lower bound for testing 3-colorability in bounded-degree graphs. In *Proceedings of the 43rd Symposium on Foundations of Computer Science*, pages 93–102, 2002.
- [Cha13] Siu On Chan. Approximation resistance from pairwise independent subgroups. In *Proceedings of the Forty-fifth Annual ACM Symposium on Theory of Computing, STOC ’13*, pages 447–456, New York, NY, USA, 2013. ACM.
- [CMM06] Eden Chlamtac, Konstantin Makarychev, and Yury Makarychev. How to play unique games using embeddings. In *Proceedings of the 47th Annual IEEE Symposium on Foundations of Computer Science*, pages 687–696, 2006.
- [CMM07] Moses Charikar, Konstantin Makarychev, and Yury Makarychev. Near-optimal algorithms for maximum constraint satisfaction problems. In *Proceedings of the 18th Annual ACM-SIAM SODA*, pages 62–68, 2007.
- [CRVW02] Michael Capalbo, Omer Reingold, Salil Vadhan, and Avi Wigderson. Randomness conductors and constant-degree lossless expanders. In *Proceedings of the 34th Annual ACM STOC*, pages 659–668. ACM, 2002.
- [DM13] Anindya De and Elchanan Mossel. Explicit optimal hardness via gaussian stability results. *TOCT*, 5(4):14, 2013.
- [GL14] Venkatesan Guruswami and Euiwoong Lee. Complexity of approximating csp with balance / hard constraints. In *Proceedings of the 5th Conference on Innovations in Theoretical Computer Science, ITCS ’14*, pages 439–448, 2014.
- [Gri01] Dima Grigoriev. Linear lower bound on degrees of positivstellensatz calculus proofs for the parity. *Theoretical Computer Science*, 259:613 – 622, 2001.
- [GSZ14] Venkatesan Guruswami, Ali Kemal Sinop, and Yuan Zhou. Constant factor lasserre integrality gaps for graph partitioning problems. *SIAM Journal on Optimization*, 24(4):1698 – 1717, 2014.
- [GUV09] Venkatesan Guruswami, Christopher Umans, and Salil Vadhan. Unbalanced expanders and randomness extractors from parvaresh–vardy codes. *J. ACM*, 56(4):20:1–20:34, July 2009.

- [GW95] Michel X. Goemans and David P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *J. ACM*, 42(6):1115–1145, November 1995.
- [Has05] Gustav Hast. Approximating max kcsdp - outperforming a random assignment with almost a linear factor. In *Proceedings of the 32Nd International Conference on Automata, Languages and Programming*, pages 956–968, 2005.
- [Kah95] Nabil Kahale. Eigenvalues and expansion of regular graphs. *J. ACM*, 42(5):1091–1106, September 1995.
- [Kho02] Subhash Khot. On the power of unique 2-prover 1-round games. In *Proceedings of the Thirty-fourth Annual ACM Symposium on Theory of Computing*, STOC '02, pages 767–775, New York, NY, USA, 2002. ACM.
- [KKMO07] Subhash Khot, Guy Kindler, Elchanan Mossel, and Ryan O'Donnell. Optimal inapproximability results for max-cut and other 2-variable csps? *SIAM J. Comput.*, 37(1):319–357, April 2007.
- [KV05] Subhash Khot and Nisheeth K. Vishnoi. The unique games conjecture, integrality gap for cut problems and embeddability of negative type metrics into l_1 . In *46th Annual IEEE Symposium on Foundations of Computer Science, 23-25 October 2005, Pittsburgh, PA, USA, Proceedings*, pages 53–62, 2005.
- [Las02] Jean B. Lasserre. An explicit equivalent positive semidefinite program for nonlinear 0-1 programs. *SIAM J. on Optimization*, 12(3):756–769, March 2002.
- [LM14] Anand Louis and Yury Makarychev. Approximation algorithms for hypergraph small set expansion and small set vertex expansion. In *APPROX/RANDOM 2014, September 4-6, 2014, Barcelona, Spain*, pages 339–355, 2014.
- [LRV13] Anand Louis, Prasad Raghavendra, and Santosh Vempala. The complexity of approximating vertex expansion. In *54th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2013, Berkeley, CA, USA*, pages 360–369, 2013.
- [MM12] Konstantin Makarychev and Yury Makarychev. Approximation algorithm for non-boolean MAX k-csp. In *APPROX/RANDOM 2012, Cambridge, MA, USA, August 15-17, 2012. Proceedings*, pages 254–265, 2012.
- [Nes00] Yurii Nesterov. Squared functional systems and optimization problems. In *High Performance Optimization*, volume 33 of *Applied Optimization*, pages 405–440. Springer US, 2000.
- [NT99] Noam Nisan and Amnon Ta-Shma. Extracting randomness: A survey and new constructions. *J. Comput. Syst. Sci.*, 58(1):148–173, 1999.
- [Par03] Pablo A. Parrilo. Semidefinite programming relaxations for semialgebraic problems. *Mathematical Programming*, 96(2):293–320, 2003.
- [Rag08] Prasad Raghavendra. Optimal algorithms and inapproximability results for every csp? In *Proceedings of the 40th Annual ACM Symposium on Theory of Computing, STOC 2008*, pages 245–254, 2008.

- [Rot13] Thomas Rothvoß. The lasserre hierarchy in approximation algorithms. <http://www.math.washington.edu/~rothvoss/lecturenotes/lasserresurvey.pdf>, 2013. Accessed: October 28, 2014.
- [RS09] Prasad Raghavendra and David Steurer. How to round any CSP. In *50th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2009, October 25-27, 2009, Atlanta, Georgia, USA*, pages 586–594, 2009.
- [RS10] Prasad Raghavendra and David Steurer. Graph expansion and the unique games conjecture. In *Proceedings of the 42nd ACM Symposium on Theory of Computing, STOC 2010, Cambridge, Massachusetts, USA*, pages 755–764, 2010.
- [RST10] Prasad Raghavendra, David Steurer, and Prasad Tetali. Approximations for the isoperimetric and spectral profile of graphs and related parameters. In *Proceedings of the 42nd ACM STOC*, pages 631–640, 2010.
- [RST12] Prasad Raghavendra, David Steurer, and Madhur Tulsiani. Reductions between expansion problems. In *Proceedings of the 27th Conference on Computational Complexity, CCC 2012, Porto, Portugal, June 26-29, 2012*, pages 64–73, 2012.
- [RT12] Prasad Raghavendra and Ning Tan. Approximating csps with global cardinality constraints using SDP hierarchies. In *Proceedings of the 23rd Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2012*, pages 373–387, 2012.
- [Sch08] Grant Schoenebeck. Linear level lasserre lower bounds for certain k-csps. In *Proceedings of the 2008 49th Annual IEEE Symposium on Foundations of Computer Science, FOCS '08*, pages 593–602, Washington, DC, USA, 2008.
- [Sha02] Ronen Shaltiel. Recent developments in explicit constructions of extractors. *Bulletin of the EATCS*, 77:67–95, 2002.
- [Sho87] N.Z. Shor. An approach to obtaining global extremums in polynomial mathematical programming problems. *Cybernetics*, 23(5):695–700, 1987.
- [Sip88] Michael Sipser. Expanders, randomness, or time versus space. *J. Comput. Syst. Sci.*, 36(3):379–383, June 1988.
- [SS96] Michael Sipser and Daniel Spielman. Expander codes. 6:1710–1722, 1996.
- [Ta-02] Amnon Ta-Shma. Almost optimal dispersers. *Combinatorica*, 22(1):123–145, 2002.
- [Tul09] Madhur Tulsiani. Csp gaps and reductions in the lasserre hierarchy. In *Proceedings of the Forty-first Annual ACM Symposium on Theory of Computing, STOC '09*, pages 303–312, New York, NY, USA, 2009. ACM.
- [TZ04] Amnon Ta-Shma and David Zuckerman. Extractor codes. *IEEE Transactions on Information Theory*, 50(12):3015–3025, 2004.
- [WZ99] Avi Wigderson and David Zuckerman. Expanders that beat the eigenvalue bound: Explicit construction and applications. *Combinatorica*, 19(1):125–138, 1999.

- [Zuc96a] David Zuckerman. On unapproximable versions of np-complete problems. *SIAM J. Comput.*, 25(6):1293–1304, 1996.
- [Zuc96b] David Zuckerman. Simulating BPP using a general weak random source. *Algorithmica*, 16(4/5):367–391, 1996.
- [Zuc07] David Zuckerman. Linear degree extractors and the inapproximability of max clique and chromatic number. *Theory of Computing*, 3(1):103–128, 2007.